

UNIVERSITE DE GENEVE
DEPARTEMENT DE BIOLOGIE
MOLECULAIRE

FACULTE DES SCIENCES
PROFESSEUR: J.H. MILLER

CARTE DE DELETIONS DETAILLEE
DU GENE LAC I D'ESCHERICHIA COLI
ET SA CORRELATION AVEC LA CARTE PHYSIQUE

T H E S E

PRESENTÉE A LA FACULTE DES SCIENCES DE L'UNIVERSITE DE GENEVE
POUR OBTENIR LE GRADE DE DOCTEUR ES SCIENCES, MENTION BIOLOGIQUE

PAR

URSULA SCHMEISSNER

DE

MEYRIN (GENEVE)

THESE NO 1865

ALLEMAGNE

BÖNECKE - DRUCK

1978



Digitized by the Internet Archive
in 2026

CARTE DE DELETIONS DETAILLEE
DU GENE LAC I D'ESCHERICHIA COLI
ET SA CORRELATION AVEC LA CARTE PHYSIQUE

T H E S E

PRESENTÉE A LA FACULTE DES SCIENCES DE L'UNIVERSITE DE GENEVE
POUR OBTENIR LE GRADE DE DOCTEUR ES SCIENCES, MENTION BIOLOGIQUE

PAR

URSULA SCHMEISSNER

DE

MEYRIN (GENEVE)



THESE NO 1865

ALLEMAGNE

BÖNECKE - DRUCK

1978

La faculté des sciences, sur le préavis de Messieurs J.H. Miller, professeur extraordinaire et directeur de thèse, L. Caro, professeur ordinaire (Département de Biologie Moléculaire) et R. Weisberg, professeur (National Institutes of Health, Bethesda, USA), autorise l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 10 mai 1978

Le Doyen :

Thèse No. 1865

Bernard LEVRAT



REMERCIEMENTS

Je tiens à exprimer mes remerciements tout d'abord au professeur J.H. Miller pour les conseils qu'il m'a donnés au cours des recherches et pour l'intérêt qu'il a porté à ce travail.

J'aimerais remercier également les professeurs L. Caro et R. Weisberg d'avoir accepté d'être mes jurés de thèse.

Ma reconnaissance s'adresse aussi à tous mes collègues de laboratoire et à toutes les personnes qui m'ont aidée dans l'accomplissement de ce travail. Je remercie en particulier B. Friedli pour sa précieuse aide technique, et C. Coulondre, A. Schmitz et H. Sommer pour leur aimable collaboration.

A toutes les personnes que j'aime

- TABLE DES MATIERES -

I) INTRODUCTION

	page
A) L'OPERON LACTOSE	1
1) Les protéines de structure de l'opéron lactose	2
2) Le promoteur lactose	4
3) L'opérateur lactose	9
B) LE REPRESSEUR LACTOSE	
1) Purification du répresseur	17
2) Structures du répresseur	18
3) Interactions du répresseur	19
4) Relation entre la structure et les fonctions du répresseur	23
C) CARTES DE DELETIONS	35
1) Formation de délétions	36
2) Construction de cartes de délétions	37
3) Pouvoir de résolution des cartes de délétions	39
4) "Effets de marqueur"	41
Figures I, II, III, IV	A

II) EXPERIENCES

A) MATERIEL ET METHODES	
1) Milieux de culture	48
2) Souches bactériennes	50
3) Sélection des délétions <u>tonB-lacI</u>	51
4) Localisation des délétions	52
5) Analyse des délétions <u>tonB</u> qui se terminent avant le gène <u>lacI</u>	56
B) CONSTRUCTION DE LA CARTE DE DELETIONS	
1) Sélection de délétions <u>tonB-lacI</u>	59
2) Distinction des différentes classes de délétions	60
3) Localisation des délétions	63
4) Corrélation de la carte génétique avec la carte physique	72

	page
5) Distribution des terminaisons de délétions dans le gène <u>lacI</u>	72
C) APPLICATIONS DE LA CARTE DE DELETIONS	
1) Fréquences de recombinaison en fonction de la distance entre délétions et mutations ponctuelles	75
2) Etude "d'effets de marqueur"	76
3) Mutations <u>I</u> ⁻ non suppressibles	78
D) DISCUSSION	
1) Carte de délétions	83
2) Distribution des terminaisons de délétions dans le gène <u>lacI</u>	84
3) Etudes de recombinaison	84
4) Distribution des mutations <u>I</u> ⁻ non suppressibles	86
ANNEXE	
E) ETUDE DU REPRESSEUR <u>LAC</u> MODIFIE PAR LA DELETION B116	
1) La séquence "leader" de l'ARN messenger <u>trp</u>	88
2) Résultats	89
3) Conclusions	90
LEGENDE DES TABLEAUX	92
LEGENDE DES FIGURES	96
TABLEAUX	103
FIGURES	121
REFERENCES	166

PREMIERE PARTIE

INTRODUCTION

A) L'OPERON LACTOSE

En 1961, l'étude de l'opéron lactose chez E.coli par Jacob et Monod a abouti à la proposition du modèle classique pour le contrôle de l'expression de gènes. Depuis lors tous les éléments impliqués dans la régulation de l'opéron lactose ont été isolés, et l'analyse détaillée de leurs interactions a contribué à en faire le système dont le mécanisme moléculaire est le mieux compris actuellement.

Ce système d'adaptation qui permet à la bactérie d'utiliser le disaccharide lactose (glucose-1,4- β -galactoside) comme seule source de carbone comprend un groupe de gènes représentant environ 0,15 % du génome d'E.coli. L'opéron lactose est composé de 3 gènes de structure adjacents, nommés Z, Y et A, qui codent respectivement pour la β -galactosidase, la lac perméase et la thiogalactoside-transacétylase. Ces gènes sont précédés par les régions de contrôle définies comme opérateur et promoteur et le gène de régulation lacI, dont le produit est le répresseur lactose. (figure I)

Le répresseur est une protéine possédant:

- un site de forte affinité pour une région spécifique de l'ADN appelée opérateur.
- 4 sites de liaison distincts pour certains sucres ou dérivés de sucres, appelés inducteurs. Les principaux sont un inducteur non-métabolisable, l'IPTG (isopropyl- β -D-thio-galactoside) et l'inducteur naturel, l'allolactose (isomère 1,6 du lactose).

En l'absence d'inducteurs, le répresseur inhibe la formation du complexe d'initiation de la transcription par son interaction avec l'opérateur. L'attachement de molécules d'inducteur sur le répresseur provoque une diminution de l'affinité du répresseur pour l'opérateur. Cet effet a pour conséquence immédiate l'induction de la synthèse d'ARN messager et une augmentation de l'expression de l'opéron par un facteur de 1000 (au maximum).

En plus d'un contrôle négatif, l'opéron lactose est soumis

à une régulation positive, par l'intermédiaire de la protéine CAP ("catabolite gene activator protein"). Cette dernière permet de répondre aux variations^{de} concentrations d'AMP cyclique intra-cellulaire. En présence d'un niveau d'AMP cyclique élevé le facteur CAP stimule d'environ 50 fois la transcription de l'opéron en s'attachant à une séquence d'ADN spécifique placée près du site de liaison de l'ARN polymérase. Le site de liaison de CAP fait donc partie du promoteur qui est défini génétiquement par des mutations affectant la transcription.

1) Les protéines de structure de l'opéron lactose

Les 3 gènes de structure désignés par Z, Y et A codent respectivement pour la β -galactosidase, la lac perméase et la thiogalactoside-transacétylase. Bien que ces gènes soient transcrits en un ARN messager polycistronique, le rapport molaire de leur taux de synthèse est d'environ 7:2:1. (Zabin, Fowler, 1970, Kennedy, 1970) Cette polarité naturelle semble due à une fréquence d'initiation de la traduction plus faible des deux derniers gènes, ce qui entraîne une dégradation préférentielle du messager peu protégé par les ribosomes (Kennell et al. 1977).

a) La β -galactosidase

C'est une protéine dont le poids moléculaire a été estimé à 540'000 (Craven et al. 1965). Elle est formée de 4 sous-unités identiques. La séquence complète des 1021 acides aminés du monomère a été déterminée récemment (Zabin et Fowler, 1972, Fowler, 1972, Fowler et Zabin, 1977).

Quand l'opéron lac est réprimé, il existe dans la cellule un "niveau basal" de β -galactosidase, d'environ 10 à 20 molécules. Après induction, une quantité d'enzyme 1000 fois plus grande est produite, et celle-ci représente environ 5 % des protéines solubles de la cellule. (Cohn, 1957)

La β -galactosidase a deux fonctions essentielles: Elle peut

- i) catalyser l'hydrolyse du lactose en ses constituants glucose et galactose,

- ii) convertir la liaison β -1,4 du lactose en β -1,6, produisant ainsi l'isomère "allolactose", qui est l'inducteur naturel de l'opéron lactose (Wallenfels et Malhotra, 1961, Jobe et Bourgeois, 1972b).

Etant donné que le lactose lui-même agit comme anti-inducteur (Jobe et Bourgeois, 1973), le niveau basal de β -galactosidase est indispensable pour amorcer l'induction naturelle.

Une mise en évidence simple de la β -galactosidase a grandement facilité l'étude de la régulation de l'opéron lactose. Elle s'effectue grâce à l'hydrolyse de β -galactosides couplés avec des substances chromogènes: l'o-nitro-phenyl- β -D-galactoside (ONPG) permet des dosages précis de l'enzyme en solution. Le 5-bromo-4chloro-3indolyl- β -D-galactoside (Xgal), un substrat non-inducteur de la β -galactosidase, offre un test sensible applicable à des cellules croissant sur milieu solide, grâce à la formation du colorant indigo. Le dosage des autres enzymes de l'opéron est plus fastidieux et n'est effectué qu'en cas de nécessité. C'est pourquoi lorsqu'il est question de "l'expression de l'opéron lactose" il s'agit, sauf indication contraire, du niveau de β -galactosidase.

b) La lac perméase

Cette protéine est également désignée par le nom "protéine M" car elle est localisée dans la membrane. Elle se lie spécifiquement à des galactosides et représente un élément essentiel dans le système de transport lac de cellules intactes. (Fox et Kennedy, 1965, Fox et al. 1967, Kennedy, 1970). Sa fonction exacte n'a cependant pas encore été démontrée in vitro.

Le poids moléculaire de la perméase est d'environ 30'000 (Jones et Kennedy, 1969, Villarejo, 1976) en présence de SDS. Il n'est pas clair si la molécule fonctionnelle est un polymère.

c) La thiogalactoside - transacétylase

La protéine produite par le gène A n'est pas indispensable au métabolisme du lactose puisque des souches sans transacétylase croissent normalement sur un milieu avec lactose. (Wilson et Kashket, 1965, Fox et al. 1966) Cette enzyme catalyse le transfert du groupe acetyl de l'acetyl-coenzyme A au groupe OH-6 d'un accepteur thiogalactoside (Zabin et al. 1959, 1962). Des expériences récentes suggèrent qu'une fonction de la transacétylase consiste à protéger la cellule contre les effets néfastes de certains analogues de lactose non métabolisables; en effet les dérivés acétylés sont excrétés. (Wilson et Kashket, 1969, Andrews et Lin, 1976) La protéine se présente sous forme de dimère. Le poids moléculaire d'une sous-unité est d'environ 32'000 (Zabin, 1963).

2) Le promoteur lactose

En général le terme promoteur désigne une région de l'ADN qui est reconnue sélectivement par l'ARN polymérase et qui permet la formation d'un complexe d'initiation pour la transcription d'un ARN messenger spécifique.

Dans l'opéron lactose (et dans divers autres opérons cataboliques) le promoteur comprend en outre un site où la protéine CAP ("catabolite gene activator protein", Riggs et al. 1971, Emmer et al. 1970) peut exercer un contrôle positif sur l'expression de l'opéron lac. La liaison de CAP, et par conséquent l'activation de la transcription, dépend directement du niveau d'AMP cyclique présent dans la cellule. (Emmer et al. 1970, Riggs et al. 1971, Anderson et al. 1971) Cette régulation permet de favoriser l'utilisation de sources de carbone plus facilement métabolisables comme le glucose, par exemple. En effet, il a été observé que le glucose provoque un abaissement du niveau d'AMP cyclique (Makman et Sutherland, 1965) et parallèlement une réduction de l'expression maximale de l'opéron lactose (Cohn et Horibata, 1959). Ce phénomène est connu sous le nom de répression cata-

bolique. (Magasanik, 1961) Deux types de mutations modifiant le niveau d'expression induit de l'opéron lactose ont été isolés:

- i) les mutations dans les gènes de structure de la protéine CAP (cap⁻, Schwartz et Beckwith, 1970) ou de l'enzyme responsable de la formation d'AMP cyclique, l'adenyl-cyclase (cya⁻, Perlman, Pastan, 1969) se distinguent par leur effet simultané sur l'opéron lactose et d'autres opérons cataboliques tels que maltose et arabinose.
- ii) les mutations touchant spécifiquement l'expression maximale de l'opéron lactose ont été localisées entre le gène lacI et l'opérateur. (Miller et al. , 1968, Ippen et al. 1968) L'étude de leurs effets a permis de subdiviser ces mutations en 3 groupes. (Beckwith, 1972, Dickson et al. 1975) Leur localisation précise par analyse de séquences a révélé l'existence de 3 régions indispensables à l'expression optimale de l'opéron lactose. Il a également été montré que 2 de ces groupes de mutations affectaient les segments d'ADN protégés par le facteur CAP et l'ARN polymérase respectivement, ce qui confirmait leur rôle.

a) Etude génétique du promoteur

Des mutations dans le promoteur d'un opéron se distinguent par les deux caractères suivants.

- Elles sont cis-dominantes, c'est-à-dire qu'elles affectent uniquement l'expression de gènes situés sur le même élément génétique.
- Elles ont un effet "pléiotropique" c'est-à-dire simultané sur tous les gènes compris dans l'opéron.

Le résultat de telles mutations peut être une diminution ou une augmentation de l'expression maximale, dans des conditions d'induction complète. Trois types de mutations ont été isolés dans le promoteur lac: (selon la classification de Dickson et al. 1975)

I Celles du premier groupe ont été interprétées comme des altérations du site d'action du facteur CAP. Des mutations telles que L8 et L29 et la délétion L1 (Scaife et Beckwith, 1966) réduisent l'expression maximale à 6 % (L8, L29) et 2 % (L1) du type sauvage. En l'absence du facteur CAP et d'AMP cyclique, dans une souche cap⁻ cya⁻, ces mutants maintiennent cependant le même niveau d'expression induite que le promoteur intact, à savoir 2 %. La délétion L1 semble donc éliminer le site d'interaction avec le facteur CAP mais laisser intacte la région permettant à l'ARN polymérase seule de promouvoir le niveau de transcription observé. (Beckwith et al. 1972).

II Une indication que l'expression résiduelle décrite ci-dessus n'était pas simplement le résultat d'une transcription commencée à un site extérieur a été fournie par d'autres mutations (classe II) situées dans le promoteur lac. L 241 et L 305 (Hopkins, 1974) réduisent le niveau induit de β -galactosidase en l'absence de CAP et d'AMP cyclique à 0,3 et 0,6 % respectivement. Par ailleurs ces mutations affectent la stimulation par CAP - AMP cyclique, qui atteint un facteur de 50 dans le cas de lacP⁺: pour L241, l'expression est stimulée 30 fois, pour L305 seulement 10 fois. Ces mutations affectent donc la transcription en l'absence de CAP. En même temps elles diminuent la capacité optimale du facteur CAP de stimuler l'ARN polymérase.

III Les mutations du 3^e type ont pour effet d'améliorer la capacité de transcription de l'ARN polymérase et de la rendre totalement ou partiellement indépendante de la présence de CAP - AMP cyclique. Les plus efficaces parmi ces mutations, appelées UV5 (Arditti et al. 1968) et p^s (de Crombrughe et al. 1971) restituent en l'absence de CAP - AMP cyclique, 50 et 20 % de l'expression maximale de l'opéron lac respectivement.

Les effets observés pour les classes II et III en l'absence de CAP et d'AMP cyclique, suggèrent que ces mutations af-

fectaient l'interaction de l'ARN polymérase avec le promoteur. Cette hypothèse a pu être confirmée, du moins en partie, par des analyses de séquences et des études de liaison in vitro, décrites ci-dessous.

b) Analyse biochimique du promoteur

La séquence de la région d'ADN comprise entre le gène lacI et l'opérateur a été déterminée par Dickson et al. (1975). Une extrémité de cette région peut être protégée par l'ARN polymérase contre l'action de l'ADNase (Gralla, non publié, voir figure II). L'analyse des séquences mutantes a révélé que les mutations du groupe III se trouvent à l'intérieur du segment protégé (Gralla, Majors, Lacy, non publié). Une autre partie de l'ADN où le facteur CAP met certaines purines à l'abri de la méthylation par le diméthylsulfate (Gilbert, 1976c) coïncide avec les mutations de la classe I. Les deux domaines protégés sont indiqués dans la figure II.

En ce qui concerne les mutations de la classe I, il est intéressant de noter que dans cette partie du promoteur, il existe deux séquences de symétrie binaire (Dickson et al. 1975, Maxam, cité par Gilbert, 1976a), et que les mutations L8 et L29 occupent des positions symétriques dans l'une d'elles (bases encadrées dans figure II) (Reznikoff, Abelson, Maxam, Sundaresan, cités par Gilbert, 1976a). Par ailleurs les expériences de méthylation (Gilbert, 1976a) ont révélé que les bases protégées étaient également disposées symétriquement dans ce fragment (indiquées en lettres grasses à la figure II). Ces résultats suggèrent que la protéine dimérique CAP forme des contacts symétriques avec l'ADN.

Deux mutations de la classe III ont été localisées dans une région d'ADN qui est protégée contre l'ADNase par la liaison avec l'ARN polymérase. Dans cette région, on observe une séquence de 7 nucléotides qui présente une bonne homologie avec les séquences analogues trouvées dans les autres promoteurs connus (Pribnow, 1975). En effet, des séquences très proches de TATAATG, connue sous le nom de "Pribnow box",

ont été retrouvées dans des promoteurs efficaces tels que λP_R , λP_L et SV40. Il est frappant de constater que les mutations UV5 (TATAATG) et p^s (TATATTG) qui rendent la transcription plus efficace, augmentent l'homologie du promoteur lac sauvage (TATGTTG) avec la séquence "idéale" TATAATG. On ne sait cependant pas si l'effet de ces deux mutations est dû à une meilleure liaison de l'ARN polymérase ou à une déstabilisation de l'ADN, qui faciliterait la dissociation des brins nécessaire pour la transcription.

Les mutations L241 et L305 de la classe II altèrent l'ADN à une distance d'environ 35 nucléotides du site d'initiation de la transcription (Dickson et al. 1977). Bien que cette région ne soit pas protégée par l'ARN polymérase, ni dans le promoteur lac, ni dans les autres promoteurs étudiés, son rôle important pour la spécificité de la transcription a été démontrée (Schaller et al. 1975). Par exemple, il a été observé avec λ et SV40 que l'ARN polymérase ne se lie spécifiquement qu'à des fragments d'ADN comprenant plus de 35 bases avant l'initiation du message. (Maurer et al. 1974, Allet et al. 1974, Meyer et al. 1975). La comparaison d'une vingtaine de séquences de promoteurs a d'ailleurs révélé que quelques bases autour de la position -35 étaient assez fidèlement conservées. Les propriétés des mutations de la classe II suggèrent que cette région participe aussi bien au fonctionnement de l'ARN polymérase que celui du facteur CAP. Toutefois, la fonction précise de cette partie d'ADN n'a pas encore été démontrée.

Bien qu'aucune évidence pour une interaction physique entre l'ARN polymérase et CAP n'ait été obtenue, ce mode d'action ne peut être exclu pour expliquer la stimulation de la transcription. En effet deux modèles ont été proposés pour les interactions sur le promoteur:

Dans le premier cas CAP entrerait en contact direct avec l'ARN polymérase par l'intermédiaire de sa sous-unité σ , (σ confère à l'ARN polymérase sa spécificité), qui serait liée à l'ADN au voisinage de la position -35. Le second modèle postule une action du facteur CAP à distance, qui produirait

un changement de l'ADN au-delà de son site de liaison pour permettre une meilleure interaction de la polymérase avec cette région. Actuellement les données expérimentales ne permettent pas de favoriser l'un ou l'autre de ces modèles.

3) L'opérateur lactose

L'opérateur a été défini par des mutations cis-dominantes provoquant une expression constitutive des enzymes lac (Jacob et Monod, 1961). Ces mutations, appelées O^c, ont été localisées entre le promoteur et le gène Z (Miller et al. 1968a). Selon ces résultats génétiques, l'opérateur a été interprété comme un site de l'ADN auquel le répresseur lac se lie pour exercer un contrôle négatif sur l'opéron lactose. Cette interaction spécifique, sensible aux inducteurs particuliers du système lac, a été démontrée in vitro. (Gilbert et Müller-Hill, 1967, Riggs et al. 1968) Il a également été montré in vitro que les mutations O^c réduisent l'affinité de l'opérateur pour le répresseur (Gilbert et Müller-Hill, 1967, Jobe et al. 1974).

La mise au point d'une technique permettant de retenir sur filtres de nitrocellulose uniquement les molécules d'ADN liées au répresseur lac (Riggs et al. 1968, 1970a) a été le point de départ d'une étude intense de la cinétique et de l'équilibre de la liaison répresseur-opérateur. (Riggs et al. 1970a, 1970b) La haute affinité entre l'opérateur et le répresseur lac ainsi que la grande spécificité de leur interaction, révélées par ces expériences, ont rendu le système lac particulièrement intéressant pour déterminer les éléments essentiels d'une liaison entre ADN et protéine.

En ce qui concerne l'opérateur lac, des indications précieuses quant à sa structure globale et ses bases indispensables ont été obtenues par divers moyens soit génétiques soit biochimiques. Les résultats de ces investigations, qui seront décrits dans les prochains chapitres, peuvent être résumés brièvement comme suit:

- i) le fragment de l'ADN protégé par la liaison du répresseur contre la digestion à l'ADNase a été isolé et sa séquence déterminée. (Gilbert et Maxam, 1973).
- ii) A l'intérieur de cette séquence, la présence du répresseur inhibe spécifiquement la modification chimique de certaines purines et de phosphates. Leur distribution dans l'ADN a révélé la configuration de l'opérateur lié au répresseur. (Gilbert, 1976b, et résultats non publiés)
- iii) L'importance de bases particulières pour la spécificité de l'interaction a été mise en valeur principalement par l'analyse des bases altérées dans des mutants O^c , (Gilbert et al. 1975) et par la comparaison avec la structure de pseudo-opérateurs. Ces derniers sont des séquences d'ADN différentes de l'opérateur, mais capables de former avec le répresseur des complexes sensibles à l'IPTG.

a) Séquence de l'opérateur

Les nucléotides numérotés 1 à 21, présentés dans la figure IIIa font partie du fragment d'ADN que le répresseur rend résistant à l'ADNase. (Gilbert et Maxam, 1973) Les bases adjacentes de part et d'autre de cette séquence ne semblent pas impliquées directement dans la liaison avec le répresseur, mais elles contribuent au caractère remarquable de cette région, à savoir une symétrie binaire extensive (bases encadrées, figure IIIa). Le rôle d'une telle symétrie dans l'interaction opérateur-répresseur a fait l'objet de plusieurs hypothèses.

b) L'interaction opérateur-répresseur

La présence d'une symétrie binaire dans la séquence de l'opérateur permettrait un réarrangement de la double hélice de l'ADN en forme de croix. Une telle structure exceptionnelle pourrait alors être stabilisée spécifiquement par le répresseur lac. Ce modèle de reconnaissance spécifique entre protéine et ADN proposé à l'origine par Gierer (1966), ne semble

cependant pas s'appliquer à l'interaction opérateur-répresseur lac. En effet la conversion de structure exigerait au minimum un déroulement de deux tours d'hélice, à savoir de 720° , et il a été montré que la liaison du répresseur provoque tout au plus une ouverture de l'ADN de 90° . (Wang et al. 1974)

Par ailleurs, une protéine peut s'attacher à l'ADN en double hélice. Dans ce cas, la reconnaissance dépend des contacts spécifiques établis entre des acides aminés de la protéine et certaines bases dans les sillons de l'acide nucléique. Ce genre d'interaction a été envisagé pour le système lac par Adler et al. (1972). Ces auteurs postulaient que les parties amino-terminales des 4 sous-unités du répresseur forment des protubérances qui s'enroulent autour de l'ADN pour établir des contacts spécifiques. La découverte de la symétrie de l'opérateur ainsi que les indications suggérant une disposition symétrique des 4 sous-unités du répresseur (Steitz, 1974) ont laissé supposer que la symétrie représente le principe de la reconnaissance spécifique.

La possibilité de protéger par le répresseur certains éléments de l'opérateur contre des modifications chimiques a permis une approche plus directe du problème. Les expériences suivantes ont été effectuées:

- i) méthylation des purines par le diméthylsulfate (DMS).
Dans le cas de la guanine, la réaction a lieu sur l'atome d'azote en position 7 qui est exposé dans le grand sillon. Pour l'adénine, il s'agit de N3 qui se trouve dans le petit sillon (Mirzabekov et Kolchinsky, 1974). Les purines protégées dans l'opérateur lac sont indiquées en lettres grasses dans la figure IIIIf. (Gilbert et al. 1976b)
- ii) la réaction des phosphates avec l'éthyl-nitro-urée (Gilbert, non publié).

L'examen d'un modèle tridimensionnel de l'opérateur a révélé que tous les éléments rendus inaccessibles aux réactifs utilisés par la présence du répresseur se trouvent d'un côté de

la double hélice. Ces expériences indiquent donc que le répresseur s'attache à la surface de l'opérateur, le long de l'axe de l'ADN. De plus, la protection de guanines et d'adénines suggère que des contacts spécifiques existent aussi bien dans le grand que dans le petit sillon. Le rôle des bases individuelles est difficile à évaluer car dans ce genre d'expériences il est à craindre que la protection soit due à un empêchement stérique plutôt qu'à une interaction directe avec le répresseur. Toutefois la distribution des purines protégées indique que le répresseur est lié à l'opérateur de façon asymétrique. L'idée d'une telle asymétrie a été confirmée lors de l'analyse de séquence des mutations O^C , contrairement à ce qui avait été proposé initialement sur la base des résultats génétiques décrits ci-après.

c) Etude des mutations de l'opérateur : O^C

Une collection d'environ 600 mutations O^C , obtenues soit spontanément soit après traitement avec 7 mutagènes différents, a été analysée en détail par Sadler et collaborateurs. (Smith et Sadler, 1971, Sadler et Smith, 1971). Les dosages des taux constitutifs de β -galactosidase dans les souches mutées ont révélé 8 classes distinctes dont l'expression atteint entre 0,5 et 20 % du niveau induit du type sauvage. Parallèlement des variations du niveau maximal de β -galactosidase (jusqu'à 50 %) ont été observées. La possibilité que ces écarts proviennent d'un effet spécifique sur le gène Z (dû à une mutation secondaire par exemple) a été éliminée par les auteurs quand bien même les contrôles montraient une stimulation préférentielle de β -galactosidase par rapport à la transacétylase dans certaines souches O^C . Malgré cette observation c'est le rapport entre les niveaux de β -galactosidase constitutif et induit (nommé valeur P) des mutants O^C qui a été choisi pour caractériser différents sites de mutations. 6 groupes (II à VII) possédant des valeurs de P distinctes ont été définis, dont chacun comprenait deux sous-classes a et b, différant jusqu'à 50 % dans leurs activités maximales de β -galactosidase. L'interprétation de ces

résultats a été guidée par des études de réversion et de mapping et probablement aussi par l'idée préconçue que la symétrie jouait un rôle essentiel dans l'interaction opérateur-répresseur. En effet Smith et Sadler ont montré que sous leurs conditions d'expérience, les différents mutagènes n'induisaient des mutations O^C que d'un facteur de 2 à 4 par rapport au background spontané. Néanmoins ces auteurs ont déduit des fréquences de révertants observées après différents traitements mutagéniques que les sous-groupes représentaient des changements de bases distincts dans l'opérateur. Cette hypothèse a été élargie à l'aide du mapping de ces mutations. Des tests de recombinaison effectués pour les groupes IIa, b, IIIa, b et IVa, b indiquaient que les membres a et b de chacun de ces groupes étaient séparés. Il aurait été intéressant de voir si ceci s'appliquait également aux groupes Va, b et VIa, b. Quant à la localisation des mutations, il ressortait clairement des résultats de croisements à quatre facteurs (effectués avec des Hfr) que des effets de marqueurs influençaient les fréquences de recombinants. Par exemple des fréquences de $2-4 \cdot 10^{-5}$ ont été observées pour les croisements entre un mutant du groupe IIa et ceux de IIIa, IVa, Vb et VIb, de même que pour IIb x IIIb, IVa et Va. D'autres croisements avec différents marqueurs extérieurs ainsi que des croisements réciproques entre diverses paires de mutations ont été effectuées pour définir l'ordre de ces groupes. Comme ces expériences n'étaient pas en mesure d'éliminer le problème des effets de marqueurs, et que certains résultats restaient ambigus, les auteurs ont tenté d'évaluer mathématiquement ces interférences, ce qui les a finalement conduits à proposer l'ordre symétrique indiqué ci-dessous. La précarité de ce mapping a été largement confirmée par l'analyse de séquence de 11 mutations appartenant, selon Sadler et Smith (1971) à différents groupes (Gilbert et al. 1975, voir ci-dessous).

Ordre proposé par Sadler et Smith:

VIIa VIb Vb IIb IIIa IVa IVb IIIb IIa Va VIa

Ordre réel:

IIa	IVa,b	IIIb	VIa,b	IIIa	IIb	VIIa	Va,b
G	G	A	G	C	G	T	C
C	C	T	C	G	C	A	G
↓	↓	↓	↓	↓	↓	↓	↓
A	T	G	T	T	A	C	T
T	A	C	A	A	T	G	A

En effet dans 3 cas (IV, V et VI) les membres a et b d'un groupe résultent du même changement de paire de bases. La fréquence de recombinants observée lors du croisement entre membres des groupes IVa et IVb reste donc sans explication, sinon celle d'une double mutation. (2 des mutants présentant un rapport anormalement élevé entre les niveaux de β -galactosidase et de transacetylase appartiennent effectivement au groupe IVa). L'analyse de séquence des mutants a d'ailleurs montré une autre faiblesse du système génétique utilisé: 3 des mutations O^C , dont 2 induites par NG (VIa,b) et 1 par 2AP (IVb) correspondent à la transversion $GC \rightarrow TA$ alors que ces mutagènes présentent une spécificité quasi exclusive pour les transitions $GC \rightleftharpoons AT$ (Yanofsky et al. 1966, Osborn et al. 1967, Müller-Hill et al. 1975). Ce fait et l'observation que certains mutants isolés dans une souche mutT (IIIa, IVb) ne proviennent pas de la transversion $AT \rightarrow CG$ (seul changement induit par mutT, Cox et Yanofsky, 1967) suggèrent que la collection de mutations O^C est composée principalement de mutations spontanées et qu'elle n'est donc pas exhaustive. Cette possibilité fournirait d'ailleurs une explication supplémentaire pour le fait étonnant qu'un seul changement de base a été détecté à chaque site. Les autres raisons possibles invoquées par Sadler et Smith (1971) à ce propos seraient que d'autres substitutions de bases n'influencent pas la liaison du répresseur ou qu'elles ont un effet létal ou très nuisible sur l'expression de l'opéron lactose.

La disposition des mutations O^C par rapport au palindrome de la séquence de l'opérateur (voir figure IIIb) est asymétrique, la majorité d'entre elles étant groupées dans la

région de gauche près du centre de symétrie. Paradoxalement deux mutations (bases 7 et 9) augmentent la symétrie du palindrome et une autre est située à son centre. Même les substitutions qui réduisent la symétrie ne soutiennent pas l'hypothèse des contacts symétriques: des mesures de liaison avec le répresseur in vitro (Jobe et al. 1974) ont révélé que les mutations situées à gauche du centre réduisent l'affinité de l'opérateur pour le répresseur beaucoup plus fortement que leurs contre-parties à droite. Par exemple le changement AT→GC des bases en position 8 résulte en une diminution de la demi-vie du complexe opérateur-répresseur de 20 fois, alors que dans la position symétrique à droite (base 14) cette même substitution ne réduit la demi-vie du complexe que d'un facteur de 3. Ces résultats suggèrent donc que la symétrie ne joue pas de rôle fondamental dans l'interaction opérateur-répresseur lac. Par conséquent on peut se demander dans quelle partie de la séquence de l'opérateur réside le principe de la spécificité. Les bases altérées par des mutations O^C sont sans doute des éléments indispensables pour le fonctionnement de l'opérateur.

Afin de définir les exigences minimales pour une liaison spécifique, on a cherché à isoler des séquences différentes de l'opérateur, mais capables de former avec le répresseur des complexes sensibles à l'IPTG. De telles séquences sont appelées pseudo-opérateurs.

d) Les pseudo-opérateurs

Le répresseur lac présente une affinité générale pour l'ADN qui est inférieure de 5 à 7 ordres de grandeur à son affinité pour l'opérateur. Toutefois de telles liaisons ne sont pas sensibles à l'IPTG. (Riggs et al. 1972, Lin et Riggs 1972, von Hippel et al. 1975, Lin et Riggs 1975a). Deux sites de liaison spécifique ont pourtant été détectés dans l'opéron lactose. L'un se trouve à l'intérieur du gène Z. Son affinité pour le répresseur est réduite d'environ 20 à 30 fois par rapport à celle de l'opérateur. In vivo, cette région ne joue qu'un rôle minime dans le contrôle de l'opéron

lactose type sauvage. Le second pseudo-opérateur, dont l'afinité correspond à peu près à celle d'un double mutant O^C dans l'opérateur lac, a été détecté à la fin du gène I. Il n'a aucun rôle physiologique.

Les séquences de ces deux régions sont indiquées dans la figure III, (lignes c et d respectivement), ainsi que les bases qu'elles ont en commun avec l'opérateur lac. (figure IIIe) La propriété la plus frappante de ces bases communes est qu'elles comprennent 6 des 8 bases altérées par des mutations O^C . Une fois de plus, leur distribution confirme l'asymétrie de l'interaction opérateur-répresseur. Néanmoins, cette constatation ne signifie pas nécessairement que la symétrie de l'opérateur est fortuite. Il est concevable que la symétrie représente le principe de reconnaissance mais que la présence de quelques bases non-symétriques affaiblit l'interaction unilatéralement de façon à créer des conditions optimales pour l'induction de l'opéron.

Le rôle des bases communes dans la spécificité de l'interaction avec le répresseur doit cependant être considéré avec précaution. En effet la proximité entre l'opéron lac et les deux pseudo-opérateurs laisse supposer que les trois séquences ont évolué à partir du même fragment. Dans ce cas le fait que certaines bases sont conservées n'indique pas nécessairement qu'elles sont indispensables pour une interaction spécifique. Aussi serait-il intéressant d'isoler d'autres pseudo-opérateurs dans des régions plus éloignées de l'opéron lactose ou sur des molécules d'ADN différentes.

B) LE REPRESSEUR LACTOSE

L'existence d'une molécule de répresseur a été proposée à l'origine par Jacob et Monod (1961) sur la base d'expériences génétiques. Celles-ci montraient que certaines mutations récessives du gène régulateur lacI résultaient en l'expression entièrement constitutive des 3 gènes de structure de l'opéron lactose.

Ce répresseur a été isolé et identifié comme une protéine. Des études in vitro ont montré que cette protéine se lie spécifiquement à l'ADN portant l'opérateur lactose. (Gilbert et Müller-Hill, 1967, Riggs et al. 1968). La mise au point de techniques très sensibles pour mesurer l'interaction entre le répresseur et l'opérateur a permis d'étudier le mécanisme de formation et de dissociation du complexe, ainsi que l'effet des inducteurs sur ces réactions.

Par ailleurs de nombreuses études génétiques et biochimiques ont été consacrées à la détermination des sites actifs du répresseur, en particulier depuis que la séquence de ses acides aminés a été élucidée. Grâce à des méthodes simples de sélection et d'analyse, des mutations affectant spécifiquement chacune des fonctions du répresseur ont été caractérisées. Une relation précise entre la séquence et les activités du répresseur a pu être établie en partie par analyse peptidique de certains répresseurs altérés, et surtout par un système génétique permettant d'établir la corrélation entre les sites de mutations (nonsense) dans le gène I et les positions correspondantes de la protéine.

Les résultats de ces travaux faciliteront l'interprétation future de la structure tridimensionnelle du répresseur qui permettra à son tour d'élucider complètement la nature des sites actifs et de leurs interactions.

1) Purification du répresseur

Le répresseur lactose ne représente que 0.003 % des pro-

téines solubles d'une cellule d'E.coli type sauvage.

Des quantités suffisantes de répresseur pour des études biochimiques peuvent être obtenues grâce à l'utilisation d'une souche portant un prophage défectif, λ dlacI^QCI₈₅₇S7, qui contient une mutation dans le promoteur du gène lacI. (Müller-Hill, 1975, Miller, 1970). L'induction de ce phage à haute température (42°) permet d'amplifier la quantité de répresseur de plusieurs centaines de fois.

Les méthodes de purification mettent à profit principalement deux propriétés du répresseur, celle de précipiter à basse concentration de sulfate d'ammonium et celle de se lier au phosphocellulose. Un inconvénient des procédés utilisés pendant longtemps était que le répresseur purifié avait perdu jusqu'à plus de 50% de son activité d'interaction avec l'opérateur (Riggs et al. 1970a). C'est pourquoi, dans toutes les expériences in vitro, où une connaissance exacte de la quantité de répresseur actif était nécessaire, des titrations par rapport à l'opérateur ont dû être effectuées.

Récemment, une méthode de purification améliorée a été proposée (Rosenberg et al. , 1977) qui semble permettre de maintenir 100 % d'activité.

2) Structures du répresseur

Le répresseur actif est une protéine composée de 4 sous-unités identiques, dont chacune possède un poids moléculaire d'environ 38000 (Müller - Hill, 1971).

La séquence du monomère déterminée par Platt et al. (1973) et Adler et al. (1972) (pour la partie amino-terminale), et par Beyreuther (1973), comprenait 347 acides aminés. Récemment il s'est avéré que cette analyse était incomplète car Farabaugh (non publié) a démontré que la séquence en nucléotides du gène lacI code en fait pour 360 acides aminés. La séquence corrigée est montrée à la figure IV.

Peu de renseignements existent sur la conformation tridi-

mensionnelle du répresseur lac. Les seuls résultats expérimentaux reproductibles proviennent de quelques études au microscope électronique. Celles-ci suggèrent une forme symétrique du tetramère, dont les dimensions ont été évaluées à $95 \times 105 \text{ \AA}$ (Brack et Humphrey, voir Bourgeois, Pfahl, 1976) ou $70 \times 80 \text{ \AA}$ (Ohshima, 1975) selon le colorant utilisé. Une analyse exacte de la structure du répresseur par diffraction de rayons X n'a pas été possible jusqu'à présent, car toutes les tentatives de cristalliser cette protéine sont restées vaines.

A défaut de données expérimentales sur le répresseur lactose Chou et al. (1975) se sont servi des relations observées chez 15 protéines à structure connue pour tenter de prédire une structure secondaire du répresseur d'après la séquence de 347 acides aminés. La validité de ce modèle est cependant très restreinte d'une part parce qu'il a été fait sur la base d'une séquence inexacte et d'autre part à cause des limitations de la méthode qui ne considère que des interactions à courte distance. Il semble en effet que la structure secondaire est déterminée également par des interactions entre des segments éloignés de la chaîne peptidique. (Sternberg et Thornton, 1978).

3) Interactions du répresseur

a) Mécanisme de l'interaction répresseur - opérateur

L'équilibre et la cinétique de la liaison du répresseur avec l'opérateur ont été étudiées dans le but d'élucider le mécanisme de répression.

Pour des mesures précises de l'interaction, une technique de rétention sur filtres de nitrocellulose (Jones et Berg, 1966) a été perfectionnée (Riggs et al. 1968). Alors que l'ADN bicaténaire libre n'est pas retenu, seul l'ADN attaché spécifiquement au répresseur lac est détecté sur de tels filtres.

La courbe de saturation obtenue en mesurant les complexes formés à l'équilibre entre une quantité fixe de λ h80dlac (marqué au ^{32}P) et des quantités croissantes de répresseur a révélé qu'un répresseur tétramérique est suffisant pour retenir une molécule d'opérateur. D'après cette expérience la constante de dissociation du complexe a été estimée à $1\text{--}2 \times 10^{-13}$ M aux conditions standard (0.01M KCL, pH 7.4) (Riggs et al. 1970). A une concentration de sel d'environ 0.2 M, considérée comme physiologique, la liaison répresseur-opérateur est plus faible et la constante d'équilibre s'approche de 10^{-11} M (Riggs et al. 1970, Lin et Riggs 1975b). Cette valeur déterminée in vitro est en accord avec la constante de 10^{-11} M proposée par Gilbert et Müller-Hill (1967) sur la base de l'induction observée in vivo, et de la concentration cellulaire du répresseur, estimée à $1\text{--}2 \times 10^{-8}$ M (10 à 20 tétramères par cellule). Le répresseur lactose a aussi une affinité générale pour l'ADN non-opérateur. Les constantes d'équilibre pour ces interactions sont cependant inférieures de quelques ordres de grandeurs à celle de la liaison répresseur-opérateur. Les valeurs obtenues par différentes techniques se situent entre 10^{-5} M et 10^{-8} M (Lin et Riggs, 1972, von Hippel et al. 1975; Wang et al. 1977).

L'analyse cinétique de la formation et de la dissociation du complexe répresseur-opérateur a également été effectuée. La réaction d'association entre le répresseur et l'opérateur (porté par λ h80dlac) s'est avérée extrêmement rapide avec une constante de vitesse de $7 \times 10^9 \text{ M}^{-1} \text{ sec}^{-1}$ (Riggs et al. 1970b). Théoriquement, la valeur maximale calculée (par Rich-ter et Eigen, 1974) en tenant compte des coefficients de diffusion des deux macromolécules ainsi que du facteur dû à l'attraction électrostatique entre le répresseur et l'opérateur (Riggs et al. 1970b) était de l'ordre de $2 \times 10^9 \text{ M}^{-1} \text{ sec}^{-1}$. Pour expliquer la différence avec la valeur expérimentale, il a été proposé que la cible effective du répresseur est plus grande que l'opérateur, (Gilbert et Müller-Hill, 1970), autrement dit que l'ADN non-opérateur agit comme "trappe" pour acheminer le répresseur plus rapidement jusqu'à l'opé-

rateur. Ceci pourrait s'effectuer soit par diffusion linéaire le long de l'ADN (Richter et Eigen, 1974) soit par transfert direct d'un site non spécifique à la séquence spécifique (von Hippel et al. 1975).

Les mesures cinétiques effectuées récemment (Goeddel et al. 1977) avec des opérateurs synthétiques comprenant 21 et 26 paires de bases ont abouti à une constante de vitesse d'association de 2.6×10^9 . Ces résultats semblent donc confirmer le rôle de l'ADN non-opérateur dans l'accélération de cette réaction.

Par opposition à la formation du complexe répresseur-opérateur sa dissociation est une réaction relativement lente, la demi-vie du complexe étant d'environ 20 min. La constante de vitesse de dissociation a été estimée à $6 \times 10^{-4} \text{ sec}^{-1}$ d'après la courbe de dissociation (Riggs et al. 1970b).

Le rapport des deux constantes de vitesse observées correspond bien à la constante d'équilibre déterminée antérieurement, vérifiant ainsi la valeur inattendue pour la réaction d'association.

$$K_{RO} = \frac{k_d}{k_a} = \frac{6 \times 10^{-4} \text{ sec}^{-1}}{7 \times 10^9 \text{ M}^{-1} \text{ sec}^{-1}} = \underline{\underline{10^{-13} \text{ M}}}$$

b) Interaction répresseur - inducteur

Tous les composés capables d'induire la synthèse de β -galactosidase in vivo affectent l'interaction entre le répresseur et l'opérateur in vitro (Riggs et al. 1970c, Barkley et al. 1975). Ces inducteurs sont des sucres ou des analogues synthétiques, en majorité des β -galactosides (Monod et al. 1951, Müller-Hill et al. 1964, Boos et al. 1967). Leur liaison au répresseur libre peut être mesurée facilement par dialyse à l'équilibre avec un inducteur marqué radioactivement (Gilbert et Müller-Hill, 1966) ou par précipitation du com-

plexe répresseur-inducteur avec du sulfate d'ammonium (Bourgeois et Jobe, 1970). Ces méthodes ont permis de déterminer qu'une molécule tétramérique de répresseur possède quatre sites de liaison pour l'inducteur et que ces sites sont identiques et indépendants (Gilbert et Müller-Hill, 1966, Riggs et al. 1970). Toutefois une étude comparative de nombreux dérivés de galactosides n'a pas permis d'élucider la nature chimique de l'interaction répresseur-inducteur. (Barkley et al. 1975) En effet, les inducteurs les plus efficaces et qui présentent la même affinité pour le répresseur ($K_d = 1,2 \times 10^{-6}$ M) sont des molécules aussi différentes que l'IPTG (isopropyl- β -D-thiogalactoside, non métabolisable par la β -galactosidase) et l'allolactose (1,6-O- β -D-galactopyranosyl-D-glucose, l'inducteur naturel, Jobe et Bourgeois, 1972a).

c) Le mécanisme de l'induction

In vivo l'addition d'un inducteur tel que l'IPTG à une culture en croissance exponentielle se traduit par une induction immédiate de la transcription de l'opéron lactose (Kepes, 1963). Etant donné la longue demi-vie du complexe répresseur-opérateur in vitro, le phénomène d'induction ne peut s'expliquer que si l'inducteur se lie à ce complexe pour en accélérer la dissociation. Il a été montré qu'en présence de 5×10^{-4} M IPTG la demi-vie du complexe n'est plus que de 5 secondes. (Barkley et al. 1975). Aux conditions d'induction complète (10^{-3} M IPTG) la constante de vitesse de dissociation a été estimée à $0,2 \text{ sec}^{-1}$. Comme la vitesse d'association entre le répresseur et l'opérateur n'est pas affectée par l'IPTG (Jobe et Bourgeois, 1972b) la constante d'équilibre de ce complexe en présence d'IPTG a pu être calculée comme suit:

$$K_{RO} (\text{IPTG}) = \frac{0,2 \text{ sec}^{-1}}{7 \times 10^9 \text{ M}^{-1} \text{ sec}^{-1}} = 3,5 \times 10^{-10} \text{ M}$$

La liaison de l'inducteur a donc pour effet de réduire l'affinité du répresseur pour l'opérateur d'un facteur d'environ

1000 (aux conditions standard de 0.01 M KCl).

In vivo le phénomène d'induction semble être couplé avec un contrôle général de la concentration de répresseur libre par l'intermédiaire des sites de basse affinité de l'ADN non-opérateur. (Lin et Riggs, 1972, Riggs et al. 1972, Lin et al. 1975, von Hippel et al. 1975). Des dosages dans des minicellules d'E.coli lac⁺ ont montré que moins de 10 % des molécules de répresseur sont libres dans le cytoplasme en l'absence comme en présence d'IPTG (Kao-Huang et al. 1977). L'interaction avec l'ADN non opérateur n'étant pas affectée par l'IPTG (von Hippel et al. 1975), la réduction de l'affinité du répresseur pour l'opérateur se traduit donc probablement par un déplacement de l'équilibre en faveur des sites de basse affinité.

4) Relation entre la structure et les fonctions du répresseur

La détermination de la séquence des acides aminés du répresseur et celle de la séquence d'ADN de l'opérateur ont suscité un intérêt particulier pour la relation structure-fonctions du répresseur. Afin de définir les régions de la protéine impliquées dans les interactions avec l'ADN de l'opérateur ou avec une molécule d'inducteur, on a tenté d'abolir ou de modifier spécifiquement une de ces activités soit par une modification chimique de la protéine soit par une modification génétique de l'ADN du gène de structure correspondant. En général les altérations de la protéine dues à la méthode génétique étaient déterminées par analyse peptidique, mais récemment un certain nombre de résidus essentiels du répresseur ont pu être identifiés simplement grâce à un système de corrélation génétique mis au point par J. Miller. Ce système sera présenté plus loin dans ce chapitre.

a) Modification chimique

Les propriétés du répresseur digéré partiellement par la trypsine ont permis de distinguer une distribution spatiale

des sites fonctionnels du répresseur. En effet, la trypsine n'enlève que les 59 résidus amino-terminaux et les 20 résidus carboxy-terminaux de la protéine native. Le "noyau tryptique" ainsi obtenu commence par le résidu 60. Il a perdu toute affinité pour l'opérateur et pour l'ADN en général mais il est resté capable de former des tetramère et de lier normalement l'IPTG (Platt et al. 1973). Ces propriétés sont également retrouvées après renaturation du "noyau tryptique" dénaturé par l'hydrochlorure de guanidine. Ces résultats montrent donc que les extrémités du répresseur ne sont pas nécessaires pour la formation de la structure tertiaire, l'aggrégation des sous-unités et la formation du site de liaison de l'IPTG.

Le rôle indispensable de la partie amino-terminale du répresseur dans l'interaction avec l'opérateur, soit que celle-ci forme des contacts directs avec l'ADN, soit qu'elle maintienne une conformation appropriée, a été confirmé par l'analyse des produits de réinitiation du répresseur lac. En effet certaines mutations nonsense au début du gène lacI provoquent une réinitiation de la traduction à partir des résidus 23, 42 ou 62 (Platt et al. 1972, Ganem et al. 1973, Files et al. 1974). Les protéines résultantes présentent des caractéristiques de structure et de fonctions semblables à celles du "noyau tryptique".

Par ailleurs certains acides aminés ont été modifiés chimiquement dans l'espoir de définir le rôle des résidus individuels. (Fanning, 1975) Cependant comme de telles modifications affectent généralement tous les acides aminés d'un type donné dans la protéine et qu'elles peuvent provoquer un empêchement stérique plutôt qu'une inactivation des résidus, l'interprétation des résultats obtenus reste limitée.

b) Modification génétique

Une corrélation plus précise entre la structure et les fonctions du répresseur lac a été obtenue grâce aux substitutions spécifiques d'un grand nombre de résidus par des

méthodes génétiques.

- le procédé classique consiste à sélectionner des mutations missense qui abolissent ou modifient une des activités du répresseur, puis à déterminer le site de l'altération par analyse peptidique de la protéine.
- Des molécules de répresseur altérées peuvent également être créées par suppression de mutations nonsense. Les mutations amber (UAG), ochre (UAA) ou opale (UGA) provoquent une terminaison précoce de la synthèse d'une chaîne peptidique (Sarabhai et al. 1964,* Brenner et al. 1967). Cet effet peut être annulé par des mutants d'E.coli, appelés supresseurs, dont certaines molécules d'ARN de transfert sont modifiées de telle façon qu'elles peuvent insérer un acide aminé spécifique à la place d'un codon nonsense. Actuellement on connaît des supresseurs insérant 6 acides aminés différents, dont la nature a été déterminée biochimiquement (Gorini, 1970).

Bien que cette méthode se limite aux sites susceptibles de se transformer en codons nonsense (il en existe 90 à 100 dans le gène lacI) la possibilité d'effectuer des substitutions multiples connues à chacun de ces sites ainsi que de corréliser les sites altérés avec les acides aminés correspondants de la protéine sans en déterminer la séquence (selon la méthode élaborée par J. Miller décrite ci-dessous) à fait de ce système un moyen d'analyse puissant.

c) Système de corrélation entre sites nonsense du gène lacI et acides aminés correspondants du répresseur.

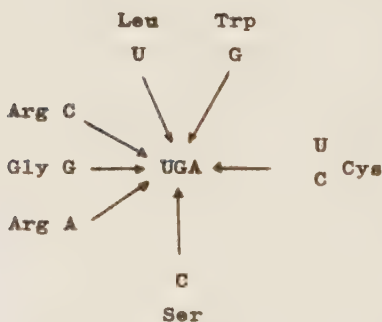
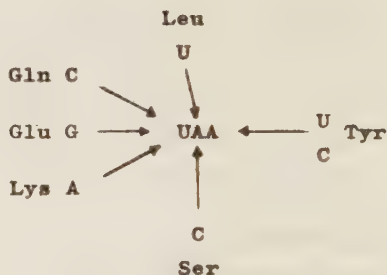
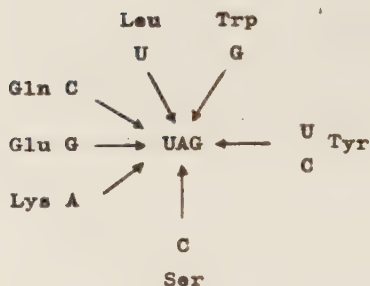
(Miller et al. 1977, Schmeissner et al. 1977a,b, Coulondre et Miller, 1977a).

Pour attribuer les sites nonsense à des positions précises dans la séquence du répresseur, différentes méthodes génétiques ont été utilisées, dont les principales seront décrites brièvement ci-dessous.

- i) Plus de 5000 mutants nonsense indépendants ont été isolés

* Brenner et Beckwith, 1965

dans le gène lacI. Ils ont été induits par plusieurs mutagènes à spécificités différentes afin de toucher le maximum des 90 à 100 sites susceptibles de se transformer en codon UAG, UAA ou UGA par les substitutions de bases indiquées ci-dessous.



- 11) Les différents sites représentés dans cette collection de mutants ont été séparés et localisés à l'aide d'une carte de délétions qui définit plus de 100 intervalles dans le gène lacI. La construction détaillée de cette carte génétique sera présentée dans la partie expérimentale. Pour distinguer deux ou plusieurs sites à l'intérieur d'un même intervalle de délétions, les caractéristiques de suppression ont été utilisées. L'effet de

différentes substitutions sur les activités du répresseur a également été testé à plusieurs températures et en présence d'inducteur pour compléter la caractérisation de l'ensemble des sites.

- iii) La corrélation des sites nonsense ainsi définis avec des positions précises du répresseur repose sur la connaissance de la séquence des acides aminés, (qui permet de prédire les résidus spécifiés par des codons "cibles") et sur celle de la spécificité des mutagènes utilisés.

Dans une première phase, les spécificités bien établies du gène mutateur mutT pour la transversion AT→CG (Cox et Yanofsky, 1967) et de la 2-aminopurine pour les transitions GC→AT (Yanofsky et al. 1966, Osborn et al. 1967) ont servi de base pour la corrélation de 31 sites nonsense.

120 mutations amber induits par mutT ont été groupées en 5 sites distincts par les moyens cités plus haut (ii). La création du codon UAG par la transversion spécifique induite par mutT est limitée à l'un des codons spécifiant la tyrosine (UAU); ces sites ont donc pu être attribués à 5 des 8 tyrosines du répresseur. Les positions exactes ont été déterminées en tenant compte du nombre d'intervalles de délétions compris entre les sites induits par mutT. Dans la partie amino-terminale du répresseur, plusieurs marqueurs déterminés par analyse peptidique ont pu être utilisés comme points de repère (Weber et al. 1975).

Les codons cibles de la 2-aminopurine sont ceux spécifiant la glutamine (CAA et CAG) et le tryptophane (UGG). L'analyse de plus de 1000 mutations nonsense indépendantes a permis de définir 14 sites amber et 12 sites ochre et de les attribuer aux deux résidus tryptophane et à 24 des 28 glutamines du répresseur.

La précision des attributions a été vérifiée pour plusieurs sites par analyse peptidique des répresseurs altérés. C'est le cas pour certains résidus (Tyr12 et 17, Gln26 et 60) de la partie amino-terminale (Weber et al. 1975) et pour

les deux tryptophanes (201 et 220, Sommer, 1977).

La concordance des résultats génétiques et biochimiques, qui confirmait d'ailleurs la spécificité des deux mutagènes 2AP et mutT, permettait donc d'utiliser les sites définis lors de cette analyse comme marqueurs pour l'attribution de nouveaux mutants nonsense.

Plus de 3500 mutants amber et ochre supplémentaires ont été isolés après induction par quatre autres mutagènes:

- l'éthyl-methanesulfonate (EMS) et la N-méthyl-N'-nitro-N-nitrosoguanidine (NG) afin de confirmer et compléter les sites obtenus par la transition GC→AT. (Osborn et al. 1967, Müller-Hill et al. 1975, Coulondre et Miller, 1977b)
- le rayonnement UV et le 4-nitroquinoline-1-oxyde pour tenter de mettre en évidence tous les sites nonsense qui peuvent être créés par les quatre transversions. (Prakash et al. , 1974, Coulondre et Miller, 1977b)

Une collection de plus de 600 mutants UGA, induits par les 6 mutagènes cités, a également été analysée.

La caractérisation des différents sites a été effectuée comme indiqué plus haut (localisation sur la carte de délétions et caractéristiques de suppression). Certains de ces sites bien définis ont pu être attribués directement à des cibles spécifiques dans des intervalles délimités par des marqueurs connus. Toutefois, pour un grand nombre d'attributions il a fallu recourir à d'autres moyens notamment:

- la spécificité de suppression, car dans certains cas, l'activité du répresseur n'est retrouvée que si l'acide aminé original est inséré.
- le nombre relativement restreint de sites UGA inductibles par les mutagènes à spécificité connue.
- la détermination de la spécificité du mutagène NQO à partir de son spectre d'induction de sites connus.

- la fréquence de recombinaison qui permet en particulier de distinguer entre mutations adjacentes (fréquences de 10^{-8} à 10^{-7}) et sites plus éloignés.

Finalement l'élucidation de la séquence en nucléotides du début de l'ARN messager du répresseur lac (Steege 1977) et de l'ADN du gène lacI (Farabaugh, 1978) a permis de confirmer tous les codons cibles et par là-même la précision des attributions. (Ces séquences ont été particulièrement précieuses pour vérifier l'attribution des 17 sites créés par substitution simultanée de deux paires de bases adjacentes).

Au total 105 sites nonsense provenant de 90 codons distincts ont été définis dans le gène lacI (dans certains cas différentes mutations nonsense peuvent être dérivées d'un même codon) et les diverses substitutions à chacun de ces sites ont créé près de 400 molécules de répresseur ayant des altérations spécifiques connues.

Cette analyse a été complétée par quelques 600 mutations missense, induites par mutT, 2AP, UV et NQO, dont les positions ont pu être déterminées avec une bonne approximation par rapport aux marqueurs nonsense précis. (Détails voir partie expérimentale). L'analyse biochimique de tous ces mutants a largement contribué à la détermination exacte des sites actifs du répresseur lac.

d) Mutations affectant les activités du répresseur

Diverses catégories de mutations du gène lacI ont été caractérisées:

- la classe la plus importante comprend les mutations I⁻ (Jacob et Monod, 1961) qui rendent le répresseur incapable de se lier à l'opérateur et conduisant par conséquent à l'expression constitutive de l'opéron lactose. On distingue d'une part les mutations I⁻ récessives à l'allèle I⁺ (dont certaines altèrent l'aggrégation des répresseurs correspondants, Schmitz et al. 1976). Quant à la plupart des muta-

tions situées au début du gène I, elles sont transdominantes (I^{-d}; Davies et Jacob, 1968), c'est-à-dire qu'elles créent des molécules capables d'inactiver le répresseur sauvage, vraisemblablement par formation de tetramères mixtes (ce phénomène a été appelé complémentation négative; Müller-Hill et al. 1968).

- Certaines mutations désignées I^s (Willson et al. 1964) provoquent une "superrépression" de l'opéron lactose qui ne peut plus être induit complètement par l'IPTG. De telles mutations sont dominantes par rapport au gène I⁺. In vitro les répresseurs I^s présentent une affinité réduite pour l'IPTG (Jobe et Bourgeois, 1972).

- Les mutations I^r (Chamness et Willson, 1970) provoquent chez le répresseur une réponse inverse à l'induction par l'IPTG, en ce sens que la répression est favorisée en présence de certaines concentrations d'IPTG. Un intérêt particulier a été suscité par la découverte que certains répresseurs I^r possèdent une affinité considérablement plus forte pour l'opérateur que le répresseur type sauvage (Jobe et Bourgeois, 1972b, Betz et Sadler, 1976, Schmitz et al. 1978).

e) Mutations affectant l'interaction répresseur-opérateur

Plusieurs analyses génétiques et physico-chimiques ont montré que les molécules de répresseur lac créées par des mutations I^{-d} entièrement constitutives sont incapables de se lier à l'opérateur mais ont maintenu leur capacité de liaison à l'IPTG in vitro (Adler et al. 1972, Weber et al. 1975). De plus la structure tetramérique de certains répresseurs I^{-d} a été démontrée in vitro (Weber et al. 1975, Schmitz et al. 1976). La dominance de tous les mutants I^{-d} est vraisemblablement le résultat d'une aggrégation entre sous-unités I^{-d} et I⁺. Par conséquent on considère que les mutations I^{-d} altèrent spécifiquement (de façon directe ou indirecte) le site d'interaction du répresseur avec l'opérateur.

Une étude portant sur plusieurs centaines de mutations missense I^{-d} , induites principalement par le mutagène NG (Adler et al. 1972, Pfahl et al. 1974) et localisées par une série de 39 délétions, suggérait que toutes les mutations I^{-d} entièrement constitutives étaient situées dans la partie du gène codant pour l'extrémité amino-terminale du répresseur. Cette localisation, limitée à la région des 60 - 70 premiers acides aminés, a été confirmée lors de l'analyse intensive de quelques 600 mutations missense (dont environ 200 I^{-d}) induites par 5 mutagènes différents (voir partie expérimentale) ainsi que des nombreux mutants obtenus par suppression de sites nonsense. Grâce à la corrélation de la carte de délétions détaillée avec la carte physique (voir partie expérimentale) il a été possible de localiser les mutations missense de façon relativement précise: elles se regroupent principalement dans deux régions, correspondant approximativement aux résidus 16 à 33 et 45 à 59. La plupart des altérations I^{-d} dont les positions précises ont été déterminées par analyse peptidique touchent les mêmes parties du répresseur (Weber et al. 1975, Müller-Hill, 1975) qui ont également été proposées comme sites d'interaction avec l'ADN de l'opérateur dans des modèles théoriques (Gursky et al. 1976, Adler et al. 1972). L'étude de substitutions multiples aux sites nonsense par les supprimeurs a par ailleurs révélé l'importance de certains résidus. En effet, aux positions 17, 18, 45, 47 et 54 seul l'acide aminé original est capable de restituer l'activité du répresseur, ce qui suggère que les résidus en question sont impliqués directement dans l'interaction avec l'ADN.

De rares mutations affectent le répresseur de façon à renforcer considérablement son interaction avec l'opérateur. La plus connue, X86, a été isolée en tant que mutation I^r (Chamness et Willson, 1970) partiellement constitutive, ayant pour effet de favoriser la répression en présence de basses concentrations d'IPTG. In vitro, le produit de cette mutation présente une affinité pour l'opérateur 50 à 100 fois plus élevée que le répresseur type sauvage (Jobe et Bourgeois, 1972b)

et il se lie aussi plus fortement à l'ADN qui ne porte pas d'opérateur lac (Pfahl, 1976). Ces effets sont causés par le changement de la serine 61 du répresseur en leucine (Files et Beyreuther, cités par Müller-Hill, 1975). Dans un autre mutant (Il2) aux propriétés semblables (J. Miller et C. Coulondre) la proline en position 3 a été remplacée par une tyrosine. La construction du double mutant X86 - Il2 a produit un répresseur dont les affinités pour l'ADN de l'opérateur et non opérateur dépassent de 1000 à 10000 fois celles du type sauvage (Schmitz et al. 1978). Des indications préliminaires que ce répresseur reconnaît spécifiquement certains sites sur l'ADN qui ne porte pas d'opérateur lac laissent espérer qu'il permettra d'isoler des séquences de pseudo-opérateurs. A partir de telles séquences, il serait peut-être possible de déduire les éléments fondamentaux de l'interaction répresseur-opérateur.

f) Mutations affectant l'interaction répresseur-inducteur

L'analyse in vitro des produits de 11 mutations I^S (Jobe et al., 1972) a montré que la majorité d'entre eux se lient normalement à l'opérateur, mais que leur affinité pour l'IPTG est fortement réduite. Ces mutations semblent donc altérer spécifiquement les parties du répresseur impliquées dans la liaison avec l'inducteur, et elles ont contribué à définir ce site.

Il faut remarquer que chez certains répresseurs I^S la capacité réduite de liaison à l'IPTG est accompagnée d'une affinité légèrement plus forte (2 à 3 fois) pour l'opérateur in vitro ou d'une altération des propriétés allostériques. De tels effets sont probablement dus à des modifications de conformation.

Afin de déterminer le site de liaison de l'inducteur, environ 150 mutations I^S, en majorité induites par la nitroso-guanidine, ont été localisées sur une carte de délétions définissant 19 intervalles dans le gène I. (Pfahl et al. 1974, Müller-Hill, 1975). Les résultats indiquaient que l'interac-

tion du répresseur avec l'inducteur pouvait être fortement affectée par certaines altération à la fin de la région amino-terminale et à quelques sites groupés dans la partie centrale de la protéine.

Cette analyse était clairement limitée à cause de la spécificité mutagénique de NG et de l'incertitude sur la longueur des intervalles de délétions. En effet, une étude plus complète de 250 mutations missense I^s, induites par plusieurs mutagènes (2AP, UV, mutT), localisées par la carte de délétions à plus de 100 intervalles et caractérisées par les niveaux d'induction obtenus en présence de différentes concentrations d'IPTG, a permis de distinguer 7 groupes correspondant à des régions limitées du répresseur. Ces régions recouvrent d'ailleurs les positions du répresseur où plusieurs substitutions d'acides aminés (par suppression de sites nonsense) réduisent fortement l'activité de liaison à l'IPTG, à savoir les positions 78, 84, 193, 220, 248, 273 et dans une moindre mesure 298 (Miller et al., non publié). De plus, des altérations de l'interaction répresseur-inducteur ont été déterminées en position 74 et 75 par analyse peptidique.

Il est difficile d'évaluer le rôle de chacun de ces sites. Toutefois comme l'étude de suppression de sites nonsense a révélé que les résidus 193, 220 et 248 étaient indispensables pour la liaison de l'inducteur, on peut penser que la structure secondaire du répresseur rapproche ces 3 positions de façon à créer un site d'interaction pour les petites molécules d'inducteur.

g) Mutations affectant l'aggrégation du répresseur

Etant donné que la structure tetramérique du répresseur est nécessaire pour une liaison efficace à l'opérateur on pouvait s'attendre à trouver parmi les mutations constitutives I⁻ certaines qui affectent spécifiquement les propriétés d'aggrégation du répresseur.

Dans ce but les mutants créés par chacun des sites

missense définis lors de l'étude de quelques 600 mutations I^- (voir partie expérimentale) ont été testés pour leur activité de liaison à l'IPTG. Outre les mutations du type I^{-d} , toutes localisées au début du gène I , deux autres groupes de mutations résultaient en l'abolition de la répression et la conservation de l'activité d'interaction avec l'IPTG. Contrairement aux I^{-d} ces mutants ne présentaient pas de transdominance. L'analyse des produits de ces mutations sur gradients de sucrose a révélé qu'ils étaient principalement sous forme de monomères avec une petite partie de dimères.

Les mutations trouvées affectent au moins 3 résidus différents: Un des sites a été localisé dans le même intervalle de délétions que celui correspondant à la tyrosine 282. Il a été montré par ailleurs que la substitution de cette tyrosine par la serine ou la glutamine (au moyen de la suppression d'un site nonsense) produit un répresseur avec des propriétés identiques à celles des mutants missense. Un autre site a été détecté à proximité du premier dans la région des résidus 278 à 298. Un groupe distinct de mutations affecte l'un des résidus 221 à 227.

C) CARTES DE DELETIONS

Dans de nombreux systèmes bactériens et phagiques ainsi que chez des organismes supérieurs comme la levure, des cartes génétiques détaillées ont contribué à la compréhension de mécanismes biologiques fondamentaux.

Pour établir une telle carte il faut localiser un grand nombre de mutations ponctuelles en les croisant entre elles et en mesurant leurs fréquences de recombinaison. Il existe cependant un procédé considérablement plus simple et plus rapide qui consiste à placer les mutations dans différents intervalles délimités par des délétions.

Une délétion est définie comme la perte d'un fragment d'ADN d'au moins quelques nucléotides. Génétiquement une telle mutation se caractérise par l'incapacité de réverter spontanément ou en réponse à un mutagène, et par l'incapacité de recombiner avec au moins deux mutations ponctuelles qui recombinent entre elles (Benzer, 1961).

Dans un croisement entre une mutation ponctuelle et une délétion, des recombinants de type sauvage ne peuvent être formés que si la mutation est située à l'extérieur de la région délétée. On peut donc définir l'ordre des mutations rapidement par une simple réponse positive ou négative. Cette technique permet d'éviter non seulement la détermination fastidieuse des fréquences de recombinaison mais encore des ambiguïtés de localisation.

En effet il a été observé dans de nombreux systèmes que les fréquences de recombinaison entre des mutations proches ne sont pas proportionnelles à la distance qui sépare ces mutations. Cette influence sur la recombinaison, qui peut être positive ou négative, provient d'un phénomène encore inexpliqué, appelé "effet de marqueur".

1) Formation de délétions

La formation spontanée de délétions est un phénomène rare dont la fréquence varie de 10^{-6} à 10^{-9} chez E.coli. Elle peut être stimulée par des mutagènes tels que HNO_2 (Tessman, 1962) et les rayonnements UV et X (Demerec, 1960). Dans le système particulier de l'opéron lactose, un facteur d'induction de 200 à 400 a été démontré pour HNO_2 et UV (Schwartz et Beckwith, 1969). Par ailleurs la présence de mutations de l'ADN polymérase I (polA) augmente la fréquence de délétions de 15 à 40 fois (Coukell et Yanofsky, 1970). On sait que ces divers facteurs peuvent provoquer des lésions dans l'ADN, soit directement, soit indirectement par interférence avec le mécanisme de réparation. Toutefois il a été montré que l'altération d'autres fonctions impliquées dans la réparation de l'ADN telles que uvrA, uvrB, uvrC et recA n'affectent pas la fréquence de délétions (Inselburg, 1967, Franklin, 1967, 1971, Anderson, 1970). L'observation que la formation de délétions ne nécessite ni une homologie extensive d'ADN ni les fonctions recA, recB et recC indique en outre que les délétions ne résultent pas d'une recombinaison générale. Le mécanisme inexpliqué a donc été appelé "recombinaison illégitime" (Campbell, 1962).

Récemment il a été montré que certains éléments de translocation tels que IS1 ou Tn10 (facteur de résistance à un antibiotique) ainsi que le phage Mu sont capables d'induire des délétions chez E.coli et Salmonella (voir revue de N. Kleckner, 1977). La présence d'une séquence IS1 dans l'opéron gal peut augmenter la fréquence de délétions des régions voisines par un facteur de 1000 (Reif et Saedler, 1975, 1977).

La formation de ces délétions présente une certaine similitude avec l'excision du prophage λ par l'intermédiaire des protéines spécifiques int et xis (Gottesman et Weisberg, 1971). En particulier toutes les délétions induites par IS1 semblent se terminer à quelques sites préférentiels dans les régions adjacentes à l'insertion (Reif et Saedler, 1977).

Il semble donc possible que les éléments d'insertion produisent également des protéines capables de diriger leur insertion et leur excision, toutefois avec une spécificité moins rigoureuse que dans le cas de λ . Un mécanisme de ce type pourrait être la cause de certains "points chauds" pour les terminaisons de délétions observés dans diverses régions d'E.coli et de Salmonella (Itikawa et Demerec, 1967, Fink et al. 1967, Spudich et al. 1970, Coukell et Yanofsky, 1971).

Il faut remarquer par ailleurs que certaines délétions créées par des excisions aberrantes du prophage λ présentent également des points de terminaison préférentiels (Pfeifer et al. 1974), alors que ce mécanisme ne semble dépendre ni de int ni de rec (Franklin, 1967, Gingery et Echols, 1968).

Par contre de nombreuses délétions (en particulier des délétions spontanées) se terminent à une multitude de sites sans spécificité apparente (Spudich et al. 1970, Hartman et al. 1971, Schleif, 1972). Cette dernière hypothèse devra être vérifiée par l'analyse des séquences aux extrémités de nombreuses délétions. Néanmoins l'ensemble des observations suggère que les délétions peuvent être formées par différents mécanismes de spécificité variable.

2) Construction de cartes de délétions

Pour pouvoir établir rapidement l'ordre de nombreuses mutations dans une région génétique donnée, il est commode de disposer d'une série de "délétions de localisation" qui commencent à une extrémité (ou à l'extérieur) et qui se terminent à différents sites à l'intérieur de cette région.

Faute d'un système de sélection approprié, de nombreuses cartes génétiques ont été construites avec des délétions trouvées dans de grandes collections de mutations spontanées ou induites. Cependant plusieurs procédés de sélection de délétions ont été mis au point, dont les principaux sont cités ci-dessous (pour plus de détails voir Beckwith, 1970 et Franklin, 1971).

- Chez le bactériophage λ , on peut sélectionner directement la perte de parties non essentielles du genome, car les délétions augmentent la résistance des phages à l'EDTA (Parkinson et Huskey, 1971). La majorité des délétions internes ainsi obtenues commencent au site att.
- L'excision aberrante du prophage λ a été utilisée pour isoler des délétions dans les parties adjacentes de l'ADN bactérien et dans certaines régions du phage (Pfeifer et Oellermann, 1967, Shapiro et Adhya, 1969). La possibilité d'insérer λ à différents sites du chromosome d'*E.coli* permet une application plus large de cette approche (Shimada et al. 1972).
- Il est possible de sélectionner l'élimination d'une région d'ADN qui est utilisée pour inhiber l'expression de certains gènes. La restauration de l'activité de ces gènes est souvent due à la fusion avec un autre gène ou opéron fonctionnel (Ames et al. 1963, Jacob et al. 1964, 1965).
- Dans l'opéron lactose des délétions ont été obtenues en sélectionnant la suppression de l'effet polaire d'une mutation ochre lacZ⁻ sur les gènes adjacents (Beckwith, 1964, Schwartz et Beckwith, 1969).
- La sélection positive de la résistance à des phages et à des colicines constitue un procédé très efficace, car il a été montré qu'une proportion importante de ces mutants résistants contiennent des délétions qui s'étendent dans les gènes adjacents. Parmi les systèmes les plus utilisés citons le site spécifiant la sensibilité aux phages T1 et ϕ 80 ainsi qu'aux colicines V et B (site tonD, Gratia, 1964) et celui des phages T3, T4 et T7 (site tfr λ , Curtiss, 1965).

Une partie importante des mutations $T1^r$ sont des délétions qui s'étendent dans les régions voisines, d'une part dans l'opéron tryptophane (Coukell et Yanofsky, 1971) et d'autre part dans le prophage $\phi 80$, lorsque celui-ci est inséré au site att80 (Franklin, 1965).

L'intérêt de ce système pour isoler des délétions dans l'opéron lactose est dû à la construction de phages transducteurs $\phi 80\Delta lac$ (Beckwith et Signer, 1966), capables d'intégrer la région lac dans les deux orientations au site att80. De cette manière de nombreuses délétions ont été isolées qui possèdent une terminaison soit dans les gènes de structure Z ou Y, et dans la région de contrôle de l'opéron (Beckwith et al. 1966, Miller et al. 1970, Beckwith, 1964, Ippen et al. 1968, Miller et al. 1968b) ou dans le gène lacI (Miller et al. 1968a, Miller et al. 1970, Ganem, 1972, Pfahl, 1972, Pfahl et al. 1974). Cette sélection a également produit un certain nombre de fusions entre l'opéron tryptophane et la région de contrôle ou les gènes de structure de l'opéron lactose (Reznikoff et al. 1969, Miller et al. 1970, Mitchell et al. 1975, 1976).

3) Pouvoir de résolution des cartes de délétions

La résolution d'une carte de délétions dépend principalement du degré de saturation obtenu avec les sites de mutations ponctuelles et du rapport entre le nombre de délétions et la longueur de la région étudiée. La séparation des sites mutationnels par les délétions est limitée par la fréquence de réversion des mutations ponctuelles et par les techniques de croisement employées.

Dans le cas du phage $T4$, la méthode de croisement par infection mixte (Tessman, 1965) permet de détecter un recombinant parmi 10^8 particules parentales. En revanche la résolution n'est que de 10^{-6} avec les procédés de localisation dans les gènes bactériens, notamment la transduction générale (Yanofsky et al. 1964) ou spécialisée (Pfeifer et

Oellermann, 1967) et les croisements entre Hfr (ou F') et F⁻ (Shapiro et Adhya, 1969, Pfahl, 1972). Toutefois on peut détecter des fréquences de recombinants au-dessous de 10^{-7} en utilisant des diploïdes qui contiennent la délétion et la mutation à tester, car les recombinants s'y accumulent pendant plusieurs générations (Ippen et al. 1968, Miller et al. 1968a,b, Ganëm, 1972).

La majorité des cartes de délétions établies jusqu'à présent définissent en moyenne des séquences de 100 nucléotides ou plus. Citons les cartes des opérons trp (Yanofsky et al. 1964, Yanofsky et al. 1971) et gal (Pfeifer et Oellermann, 1967, Pfeifer et Rosypal, 1969, Shapiro et Adhya, 1969), et celle du gène lacZ (Zipser et al. 1970).

Une meilleure résolution a été obtenue dans la région rII du phage T4 où les intervalles de délétions ont une longueur moyenne de 75 nucléotides (Benzer, 1961). Tel est aussi le cas dans les gènes araB et araC de l'opéron arabinose Schleif, 1972, ainsi que dans le gène lacI (Pfahl et al. 1974) où des segments d'environ 50 nucléotides ont été définis. Une carte particulièrement détaillée avec des intervalles de 10 nucléotides en moyenne a été obtenue pour le gène de l'iso-1-cytochrome c de la levure (Sherman et al. 1975).

Si de telles délétions permettent d'établir les positions relatives des sites de mutations, elles ne donnent par contre aucune indication sur leurs positions exactes.

Une corrélation approximative entre la carte génétique et la carte physique peut être obtenue en mesurant la longueur des fragments de protéine créés par des mutations nonsense (Sarabhai et al. 1964). Cette méthode a été utilisée pour 30 mutations dans rIIA et 15 dans rIIB (A. Bolle et R. Epstein, communication personnelle).

Si la séquence des acides aminés de la protéine est connue, on peut établir une relation plus précise en déterminant par analyse peptidique les altérations dues aux différentes mutations. Une telle corrélation a été obtenue pour 15 mutations dans le gène de la tryptophane-synthétase (280

acides aminés) et pour 15 mutations dans le gène de l'iso-1-cytochrome c (109 acides aminés) chez la levure. (Carlton et Yanofsky, 1963, 1965, Guest et Yanofsky, 1965, Moore et Sherman, 1977) Le même procédé a été utilisé pour 11 mutations du gène lacI groupées dans la région amino-terminale (61 acides aminés) du répresseur lactose (Weber et al. 1975). C'est pour étendre cette caractérisation à l'ensemble du gène lacI que nous avons entrepris d'établir la corrélation entre sites de mutations et acides aminés correspondants par des moyens génétiques (voir chapitre B) 4c de l'introduction et Miller et al. 1977, Coulondre et al. 1977a). Un de ces moyens est une carte de délétions détaillée du gène lacI qui sera décrite dans la partie expérimentale.

4) "Effets de marqueur"

Les cartes de délétions ont été largement utilisées à une grande échelle, par exemple pour confirmer l'ordre des gènes dans certains opérons ou l'insertion linéaire d'un phage dans le chromosome bactérien. Toutefois c'est pour la détermination de la structure fine d'un gène, c'est-à-dire pour la localisation d'un grand nombre de mutations dans une région génétique relativement restreinte qu'elles ont une importance particulière.

Il a été observé dans plusieurs microorganismes que la recombinaison entre des mutations ponctuelles proches est influencée par des "effets de marqueur". (voir revue par Stadler, 1973) Ces derniers perturbent l'additivité des fréquences de recombinaison et conduisent souvent à des résultats ambigus concernant l'ordre des mutations.

Les causes exactes des "effets de marqueur" ne sont pas encore comprises. Dans de nombreux cas les fréquences de recombinaison semblent être déterminées par la substitution spécifique au site de mutation et/ou les nucléotides voisins.

Chez certains phages des effets similaires peuvent être provoqués par des "points chauds de recombinaison" qui

stimulent la recombinaison pour des régions limitées. Un site de ce genre a été trouvé chez le phage T4 de type sauvage (Beckendorf et Wilson, 1972), et chez le phage λ , plusieurs "points chauds" ont été créés par mutation (Lam et al. 1964).

Des "effets de marqueur" spécifiques sur la recombinaison entre nucléotides adjacents dans la région rII du phage T4 ont pu être mesurés grâce à la méthode de Tessman (1965). Ronen et Salts (1971) ont croisé des paires de mutations UAG et UGA dérivées du même codon, à 12 sites différents et ils ont observé des variations jusqu'à un facteur de 1000 dans les fréquences de recombinants.

Pourtant les variations dues aux "effets de marqueur" ne se limitent pas à des mutations adjacentes.

Dans le gène trpA différents croisements entre 15 mutations à des positions connues ont révélé des "effets de marqueur" pour des distances comprises entre 30 et plus de 600 nucléotides (Stadler et Karija, 1973). Les fréquences de recombinants mesurées pour diverses mutations dans le gène lacZ dévient aussi fortement d'une fonction linéaire jusqu'à des distances de 2000 à 3000 paires de bases (Norkin, 1970). Lors de l'étude détaillée du gène lacI, Coulondre et Miller (1977a) ont observé des fréquences de recombinants différant d'un facteur de 1000 pour des paires de mutations séparées par 10 nucléotides, situées à divers endroits du gène.

Il a également été montré que diverses substitutions de la même base (ou de bases adjacentes) peuvent avoir des influences contraires sur la recombinaison avec une autre mutation. Un exemple frappant de ce phénomène (dans le gène trpA) est l'observation que les deux dérivés UAG et CAG du même codon GAG recombinent avec une autre mutation à des fréquences différant d'un facteur de 15 (Drapeau et al. 1968). Un effet similaire a été observé pour des mutations dans le gène de l'iso-1-cytochrome de la levure (Moore et Sherman, 1977): Le dérivé CUG du codon initial AUG présente des fréquences de recombinaison jusqu'à 7 fois plus élevées que le

dérivé GUG dans des croisements avec plusieurs mutations à des distances jusqu'à 15 paires de bases.

Ces deux exemples suggèrent que l'appariement des bases C:T favorise la recombinaison par rapport à C:C ou G:T. Il faut cependant attendre les résultats d'autres études de ce genre pour tirer des conclusions quant à la spécificité des "effets de marqueur".

L'OPéron LACTOSE

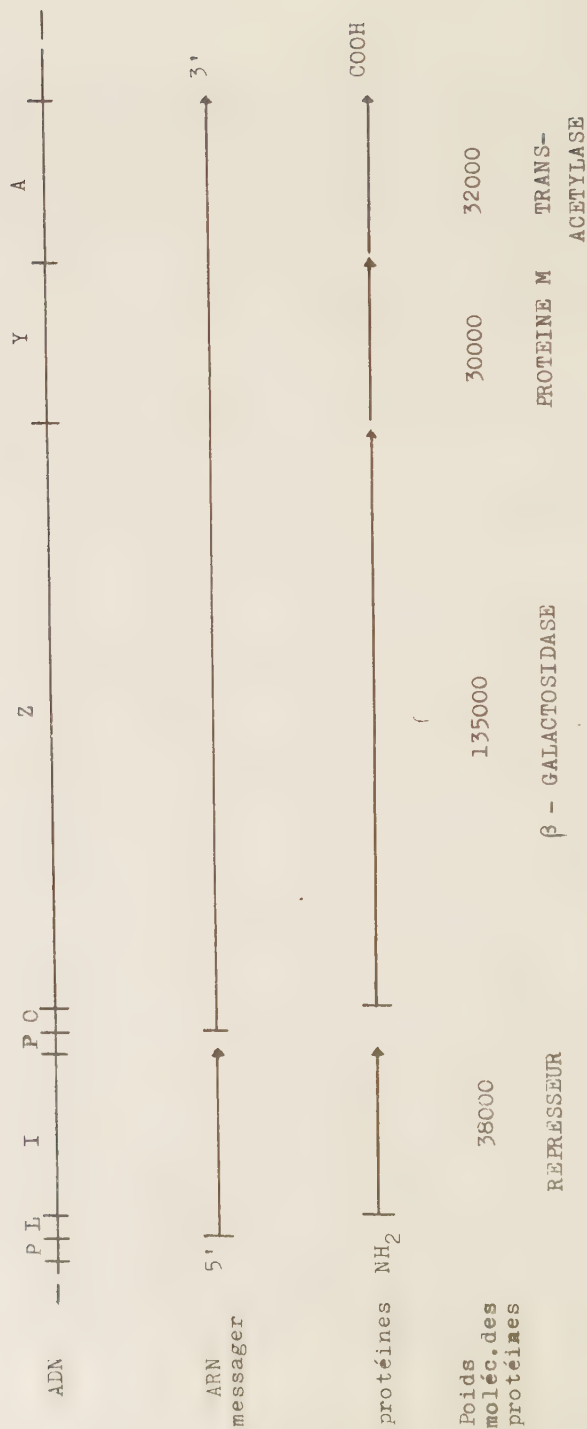
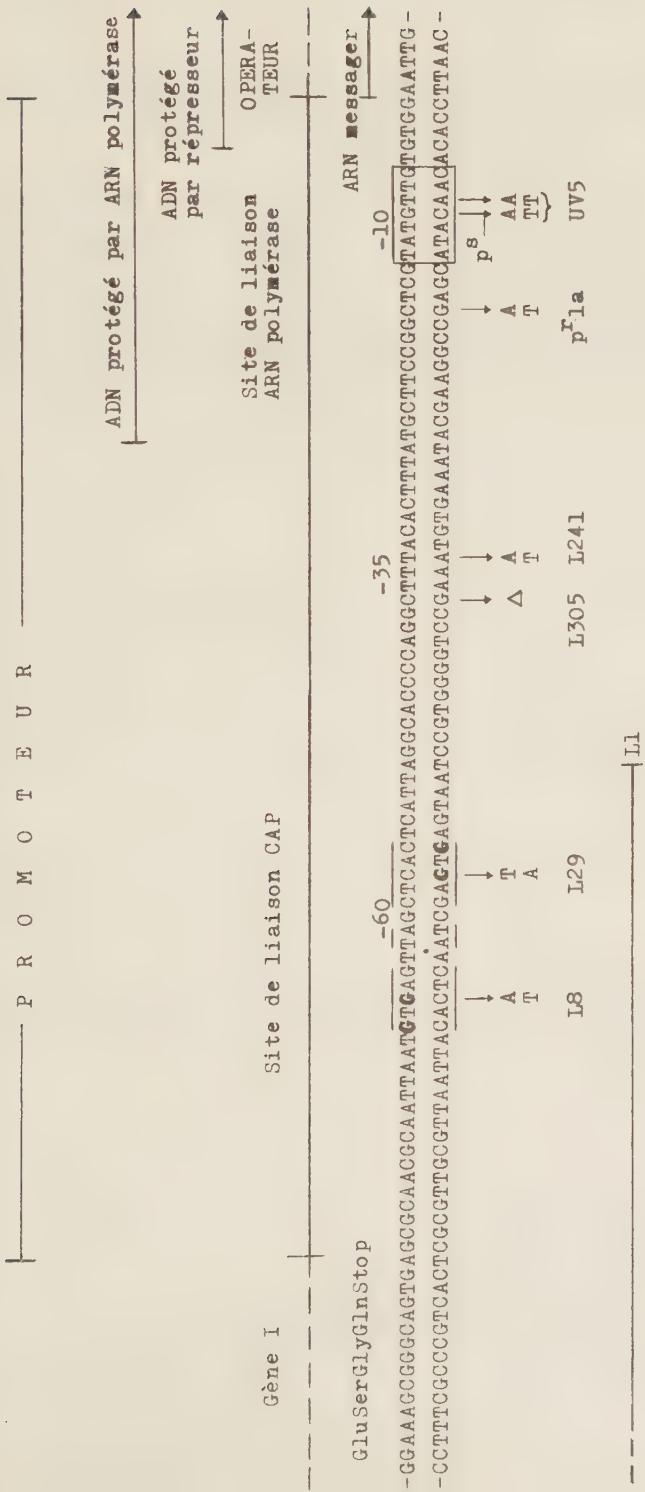


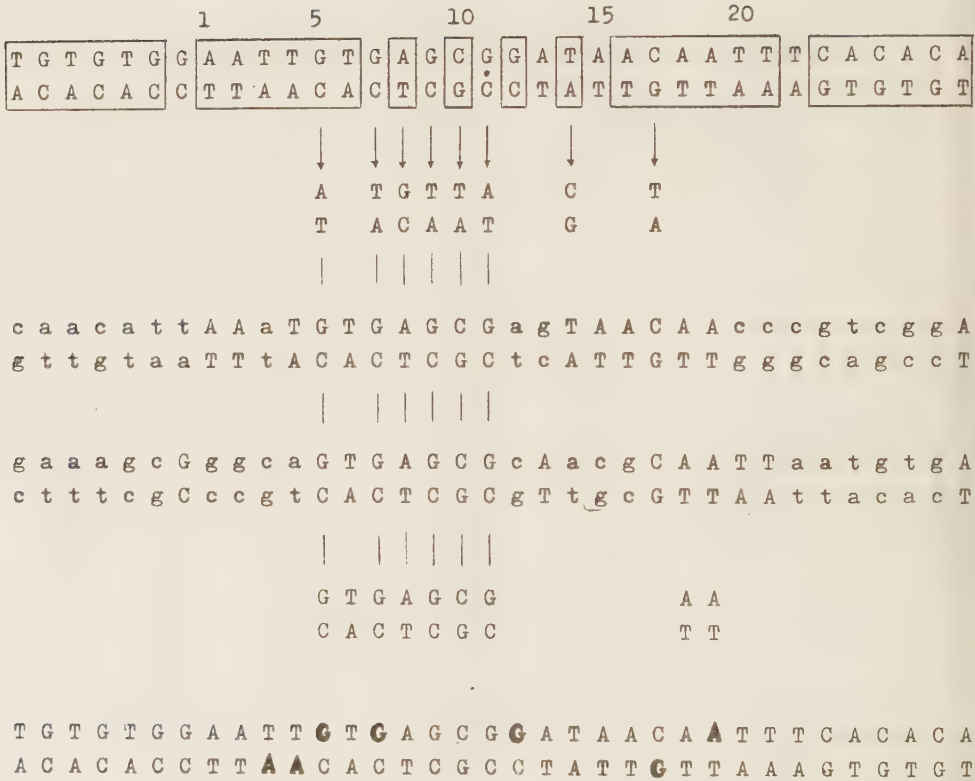
Figure I



G: bases protégées par CAP

Figure II

L'OPERATEUR LACTOSE



- a: Opérateur lac
- b: Mutations O^c
- c: Pseudo-opérateur du gène lacZ
- d: Pseudo-opérateur du gène lacI
- e: Bases communes aux deux pseudo-opérateurs
- f: Lettres grasses: Méthylation inhibée par répresseur

Figure III

LA SEQUENCE DES ACIDES AMINES DU REPRESSEUR LACTOSE

NH ₂ -	
met-lys-pro-val-thr-leu-tyr-asp-val-ala-glu-tyr-ala-gly-val-	15
ser-tyr-gln-thr-val-ser-arg-val-val-asn-gln-ala-ser-his-val-	30
ser-ala-lys-thr-arg-glu-lys-val-glu-ala-ala-met-ala-glu-leu-	45
asn-tyr-ile-pro-asn-arg-val-ala-gln-gln-leu-ala-gly-lys-gln-	60
ser-leu-leu-ile-gly-val-ala-thr-ser-ser-leu-ala-leu-his-ala-	75
pro-ser-gln-ile-val-ala-ala-ile-lys-ser-arg-ala-asp-gln-leu-	90
gly-ala-ser-val-val-val-ser-met-val-glu-arg-ser-gly-val-glu-	105
ala-cys-lys-ala-ala-val-his-asn-leu-leu-ala-gln-arg-val-ser-	120
gly-leu-ile-ile-asn-tyr-pro-leu-asp-asp-gln-asp-ala-ile-ala-	135
val-glu-ala-ala-cys-thr-asn-val-pro-ala-leu-phe- <u>leu-asp-val-</u>	150
<u>ser-asp-gln-thr-pro-ile-asn-ser-ile-ile-phe-ser-his-glu-asp-</u>	165
gly-thr-arg-leu-gly-val-glu-his-leu-val-ala-leu-gly-his-gln-	180
gln-ile-ala-leu-leu-ala-gly-pro-leu-ser-ser-val-ser-ala-arg-	195
leu-arg-leu-ala-gly-trp-his-lys-tyr-leu-thr-arg-asn-gln-ile-	210
gln-pro-ile-ala-glu-arg-glu-gly-asp-trp-ser-ala-met-ser-gly-	225
phe-gln-gln-thr-met- <u>gln-met</u> -leu-asn-glu-gly-ile-val-pro-thr-	240
ala-met-leu-val-ala-asn-asp-gln-met-ala-leu-gly-ala-met-arg-	255
ala-ile-thr-glu-ser-gly-leu-arg-val-gly-ala-asp-ile-ser-val-	270
val-gly-tyr-asp-asp-thr-glu-asp-ser-ser-cys-tyr-ile-pro-pro-	285
leu-thr-thr-ile-lys-gln-asp-phe-arg-leu-leu-gly-gln-thr-ser-	300
val-asp-arg-leu-leu-gln-leu-ser-gln-gly-gln-ala-val-lys-gly-	315
asn-gln-leu-leu-pro-val-ser-leu-val-lys-arg-lys-thr-thr-leu-	330
ala-pro-asn-thr-gln-thr-ala-ser-pro-arg-ala-leu-ala-asp-ser-	345
leu-met-gln-leu-ala-arg-gln-val-ser-arg-leu-glu-ser-gly-gln-	360
	COOH

Figure IV

Séquence déterminée par Beyreuther et al.(1973), complétée par les acides aminés (soulignés) déduits de la séquence de nucléotides du gène lacI (P. Farabaugh, non publié)

SECONDE PARTIE

EXPERIENCES

A) MATERIEL ET METHODES

1) MILIEUX DE CULTURE

Pour la composition et la préparation des milieux de culture courants il est recommandé de consulter le manuel de Miller (1972).

a) Milieux riches:

- liquides : LB, 2YT
- solides (pour boîtes de Petri): LB, H(pour lysats de phages) et agar H-top.

b) Milieux minima

Tous les milieux de culture minima utilisés lors de ce travail sont à base de milieu A. Celui-ci contient par litre 10.5 g de K_2HPO_4 , 4.5 g de KH_2PO_4 , 1 g de $(NH_4)_2SO_4$ et 1 g de citrate de sodium. $2H_2O$. (La présence de citrate est essentielle pour la croissance de souches tonB sur milieu solide, car ces mutants sont très sensibles aux traces de Cr^{+++} contenues dans l'agar commercial.) Ce milieu a généralement été complété de $MgSO_4$ à 1 mM, de thiamine (vitamine B1) à 8 $\mu g/ml$ et de sources de carbone telles que glucose ou arabinose à 0,4 % de concentrations finales. En cas de besoin, les acides aminés (ou analogues) et antibiotiques ont été ajoutés aux concentrations suivantes:

tryptophane, proline		80 $\mu g/ml$
5-methyltryptophane	(Aldrich)	100 $\mu g/ml$
streptomycine	(Pfizer)	200 $\mu g/ml$
acide nalidixique	(Sigma)	40 $\mu g/ml$
rifampicine	(Sigma)	100 $\mu g/ml$

L'agar F-top a été employé pour l'étalement de mélanges d'infection sur milieu minimal.

c) Boîtes de sélection spéciales

Pgal: Pour la sélection de mutants I⁻ (et de P⁺ dans le cas particulier du mapping contre la mutation L8 du pro-

moteur lac, décrit plus loin), le phényl- β -D-galactoside (Pgal, Sigma) est utilisé comme seule source de carbone, à une concentration de 750 μ g/ml. Ce composé ne peut induire l'opéron lactose.

Pgal-glucose: Milieu minimal glucose avec 500 μ g/ml de Pgal pour la sélection de colonies I^+ (dans un back-ground génétique gal E).

TONPG: Milieu minimal contenant du succinate de sodium à une concentration finale de 0.2 % et 300 mg/l d'ortho-nitro-phenyl- β -D-thiogalactoside (TONPG, Bachem) ajoutés en solution dans 2.5 ml de N,N-diméthylformamide. Le TONPG qui n'est pas un inducteur de l'opéron lac inhibe la croissance de mutants avec un niveau constitutif de lac permease (produit du gène lac Y.)

d) Boîtes indicatrices

McConkey Lac: Sur ce milieu les colonies lac⁺ sont rouge foncé, les colonies lac⁻ blanches. On peut également distinguer les colonies lac P⁻ (rouge clair) des lac P⁺ (rouge foncé)

Xgal: Milieu minimal glucose avec 40 mg/l de 5,bromo-4, chloro-3, indolyl- β -D-galactoside (Xgal, Bachem) qui sont ajoutés dissous dans 1 ml de N,N,diméthylformamide. Xgal est un substrat, mais non pas un inducteur de la β -galactosidase. Un des produits d'hydrolyse est le 5,bromo-4, chloro-indigo, de couleur bleu foncé. Par conséquent, les colonies I^- apparaissent bleu foncé sur ce milieu alors que les colonies I^+ ont une coloration bleu pâle.

e) Tampons

- Z : pour le dosage de la β -galactosidase
- TMS : pour la préparation d'extraits de cellules et pour la dialyse.

Nom	Sexe	Marqueurs génétiques	Remarques et Références
X7800	F ⁻	<u>ara val^r Δ(lacproB)</u> <u>galE</u> <u>strA</u> <u>thi</u> (<u>ø80dlac⁺</u>)	<u>ø80dlac</u> : Beckwith, Signer, 1966 <u>Δ(lacproB)</u> X111: Cuzin, Jacob, 1964 <u>galE</u> FL132: Malamy, 1966
X7801,2..F ⁻		Dérivés de X7800 avec délétions <u>tonB</u> qui s'étendent dans le gène <u>lacI</u> sur <u>ø80dlac</u>	
X8726	F ⁻	Souches de délétions semblables à la série	Délétions isolées par J. Miller
X8730		X7801,2..., mais dérivées de X8600	(Miller et al. 1970)
X7026r	F ⁻	<u>Δ(lacproB)</u> <u>supE</u> <u>recA</u> <u>thi</u>	<u>supE</u> : supprimeur de mutations amber insérant la glutamine
S90CR	F ⁻	<u>ara Δ(lacproB)</u> <u>strA</u> <u>rif</u> <u>thi</u>	
P90CN	F ⁻	<u>ara Δ(lacproB)</u> <u>nalA</u> <u>thi</u>	
CSH 52	F ⁻	<u>ara Δ(lacproB)</u> <u>recA</u> <u>strA</u> <u>thi</u> (<u>ø80dlac⁺</u>)	
X7733	F ⁻	<u>Δ(lacproB)</u> <u>galE</u> <u>strA</u> <u>thi</u>	
X7734	F ⁻	<u>Δ(lacproB)</u> <u>galE</u> <u>recA</u> <u>strA</u> <u>thi</u>	
CSH 61	Hfr	<u>trpR</u> <u>thi</u>	<u>trpR</u> : Imamoto et al. 1966
GM1	F ⁻	<u>F⁻lacpro(I⁰I8)</u> / <u>ara Δ(lacproB)</u> <u>thi</u>	F ⁻ <u>lacproB</u> isolé par E.R. Signer I ⁰ : Müller-Hill et al., 1968 I8 : Scaife, Beckwith, 1966
CSH 48	F ⁻	<u>colV_B</u> <u>trp</u> <u>cys</u> / <u>his</u> (<u>ø80</u>) <u>thi</u>	Episode de Frédéricq, 1963 (voir Signer et al, 1965)
CSH 33	F ⁻	<u>colV_B</u> <u>trp</u> <u>cys</u> / <u>thi</u>	

3) Sélection des délétions ton B - lac I

a) Préparation du lysat de colicines

Ces lysats ont été obtenus par induction à la lumière UV de la souche CSH 48, lysogène pour $\phi 80$ et les colicines V et B, comme décrit par Gottesman et Beckwith (1969).

b) Préparation du lysat de $\phi 80$ vir (voir Miller, 1972)

Les stocks de phages $\phi 80$ vir ont été préparés par infection de 10^8 bactéries avec $1-2 \times 10^5$ phages $\phi 80$ vir sur boîtes de Petri (milieu H). Les titres obtenus étaient généralement de 2×10^{11} centres infectieux par millilitre.

c) Sélection de délétions ton B qui affectent le gène lac I

La méthode de sélection décrite par Gottesman et Beckwith (1969) a été adaptée de façon à donner un rendement optimal de délétions ton B - lac I comme suit: Des cultures de 2.5 ml LB, obtenues à partir de colonies indépendantes de la souche X7800, ont été incubées avec aération à 37°C pendant la nuit. Le titre atteignait généralement 4 à 5×10^9 cellules/ml. Après centrifugation et écartement du surnageant, le culot a été resuspendu dans les quelques gouttes de liquide restant. 0.3 ml de lysat de colicines V et B et 0.1 ml de lysat de $\phi 80$ vir (multiplicité d'infection de 2 par bactérie) ont été ajoutés. Les mélanges ont été incubés à 37°C pendant 30 min. pour la préadsorption, puis répartis, avec 4 ml de F-top agar fondu (50°C) sur des boîtes à milieu minimal glucose avec tryptophane, proline et Xgal.

Après 24 - 30 heures d'incubation à 37°C , 500 - 1000 survivants apparaissaient par boîte. Dans 86 % des cas, une ou deux colonies lac constitutives pouvaient être distinguées grâce à leur coloration bleue. Un seul mutant I⁻ de chaque culture a été gardé. Par cette méthode 337 colonies lac constitutives ont été isolées à partir de 390 cultures.

Cependant une partie de la sélection a été faite dans des

conditions moins favorables: L'addition de 0.3 à 0.4 ml de lysat de Ø 80 vir, qui représente une multiplicité d'infection de 6 à 8 par bactérie, diminuait le nombre de survivants. Ce procédé n'a permis d'obtenir que 68 colonies I⁻ à partir de 150 cultures. L'utilisation combinée de 2YT comme milieu de culture et de 0.4 ml de lysat de Ø 80 vir a réduit le rendement de façon encore plus drastique. Dans ces conditions, seulement 2 colonies I⁻ ont été trouvées dans 50 cultures.

4) Localisation des délétions

a) Construction de diploïdes (Miller, 1972)

Tous les mutants I⁻ utilisés lors de ce travail ont été isolés sur un épisome F'lac pro et conservés dans une souche Δ(lacproB), rec A, str^s (X7026r). Les souches réceptrices porteuses de délétions ton B - lac I (X7800) sont également délétées de la région lac pro mais résistantes à la streptomycine. Les diploïdes désirés ont donc été obtenus par sélection du marqueur pro⁺ et contre-sélection à l'aide de la streptomycine. Pour le transfert de l'épisome F'lac pro dans des souches telles que S90CR ou P90CN, la rifampicine et l'acide nalidixique ont été utilisés comme agents de contre-sélection respectifs.

Les croisements ont été effectués de deux manières:

- Par répllication sur milieu solide

ce qui permet le croisement simultané d'un grand nombre de mutants. Le soir environ 50 colonies F'lac ont été disposées en grille sur une boîte LB et le lendemain matin celle-ci a été répliquée sur plusieurs boîtes LB selon le nombre de souches réceptrices à tester. Après 6 à 8 heures d'incubation à 37°C, chacune de ces boîtes a été répliquée sur un tapis d'une souche F⁻ différente. Ces tapis avaient été préparés préalablement en étalant 0.2 ml de cultures saturées de souches F⁻ sur les boîtes de sélection adéquates. Ces boîtes ont été incubées à 37°C pendant la nuit.

- En milieu liquide

Pour les croisements individuels, 0.2 ml d'une culture en croissance exponentielle de la souche F'lac et 0,2ml d'une culture saturée de la souche réceptrice ont été mélangés et placés au bain-marie à 37°C pendant environ 1 heure. Ensuite une goutte de ce mélange a été répartie sur le milieu de sélection approprié à l'aide d'une anse de platine de façon à obtenir des colonies isolées des diploïdes désirés.

b) Méthode de localisation par réplication

Le principe de la méthode qui vise à éliminer tous les mutants à l'exception des recombinants I^+ est expliqué en détails au chapitre des résultats. Pratiquement cette technique consiste en 3 réplications successives: Tout d'abord les croisements entre les mutants I^- et les souches de délétions ont été effectués par réplication comme décrit plus haut. Puis les boîtes portant les quelques 50 blocs de diploïdes ont été répliquées sur milieu minimal contenant du succinate, tryptophane et TONPG, afin d'éliminer les colonies avec un niveau constitutif de lac perméase. Ces boîtes TONPG ont été placées à 39.5°C, (une température qui semble retarder ou réduire considérablement l'apparition d'un background et de contaminants indésirables) et répliquées, 36 à 48 heures plus tard, sur milieu minimal glucose additionné de tryptophane et de Pgal (500 µg/ml). Grâce à la mutation gal E, présente dans toutes les souches de délétions, les mutants I^- lysent sur ce milieu, alors que les colonies I^+ survivent. Les résultats pouvaient généralement être lus après 15 heures d'incubation à 37°C, comme le montrent les exemples représentés dans la figure 3. Dans certains cas, des incubation plus longues, de 24 à 30 heures, à une température de 39.5°C, ont été nécessaires pour une bonne résolution.

c) Méthode de localisation à haute résolution (D. Ganem, 1972)

Ce procédé repose également sur la sélection de recombinants

I⁺ (dans un background génétique galE) par Pgal, il permet en outre de déterminer leur fréquence exacte. (Pour détails théoriques voir partie expérimentale). Des colonies isolées de "diploïdes de localisation" ont été inoculées dans 2.5ml de milieu 2YT et cultivées pendant la nuit. La densité de saturation atteignait généralement 4 à 5×10^9 cellules par millilitre. 0.05ml d'une dilution 1:5 de ces cultures, c'est-à-dire 4 à 5×10^7 cellules, ont été répartis sur des boîtes à milieu minimal glucose + tryptophane + Pgal. Celles-ci ont été incubées à 39.5°C pendant 36 heures, puis les survivants de la sélection Pgal ont été répliqués sur milieu indicateur Xgal. Après environ 15 heures d'incubation à 37°C, les différents types de survivants ont pu être distingués par la coloration des colonies:

- bleu foncé: les révertants de la mutation galE
- blanches: les homogénètes Z⁻/Z⁻
- bleu pâle: les recombinants I⁺

2 à 10 colonies pâles ont été vérifiées dans chaque cas pour leur caractère inductible sur milieu indicateur Xgal sans et avec IPTG (10^{-3} M).

Un petit test supplémentaire a dû être effectué pour quelques souches porteuses de délétions tonB qui se terminent dans la région de contrôle du gène lacI. Dans certaines de ces souches l'expression du gène lacI a été mise sous contrôle trp (détails voir distinction des délétions). Des mutations spontanées conduisant à l'expression constitutive de l'opéron trp peuvent donc créer un phénotype I⁺ et interférer avec les résultats de recombinaison. Afin de tester si un gène I⁺ avait été reconstitué par recombinaison sur l'épisome, ces derniers ont été transférés dans une souche Δlacpro (S90CR) par répllication, puis leur caractère I⁺ a été vérifié sur milieu indicateur Xgal sans et avec IPTG.

d) Test de réversion

Dans plusieurs croisements un très petit nombre de

colonies I^+ a été observé. Dans ces cas les fréquences de recombinants ont été analysées plus précisément en examinant 6 à 10 fois plus de cellules que normalement. Parallèlement les fréquences de révertants ont été déterminées dans les diploïdes adéquats (où les délétions couvrent les mutations ponctuelles en question). En général 3 à 5 colonies indépendantes d'un diploïde ont été inoculées respectivement dans 5 ml LB, puis 0.5 ml de chaque culture saturée ont été sous-cultivés deux fois successivement. Ensuite environ 10^8 cellules par culture ont été étalées sur des boîtes à milieu minimal-glucose-tryptophane-Pgal, selon le procédé de localisation décrit plus haut. Toutes les colonies apparaissant blanches ou pâles sur le milieu indicateur Xgal ont été testées pour leur inductibilité. Si un diploïde produisait dix fois plus de colonies I^+ que le contrôle, le résultat de recombinaison était considéré comme positif.

e) Localisation de délétions par rapport à la mutation L8
du promoteur lac.

Les délétions qui se terminent au-delà des marqueurs connus du gène lac I ont pu être placées de façon simple par rapport à la mutation L8 du promoteur lac grâce à un pseudo-révertant d'une mutation nonsense au début du gène I. Ce pseudo-révertant qui a été isolé sur un épisome F'lac pro porteur de L8 (lors d'une étude secondaire de ce travail) provoque une expression faiblement constitutive de β -galactosidase. Alors que cette dernière ne permet pas à la souche de croître sur un milieu Pgal, le niveau d'enzyme produit en présence de l'allèle P⁺ est suffisant pour la croissance sur ce milieu. Ceci qui constitue un excellent moyen de sélectionner des recombinants P⁺.

Les diploïdes contenant le pseudo-révertant et la mutation L8 sur l'épisome et les délétions à tester ainsi que l'allèle P⁺ sur le chromosome, ont été cultivés pendant la nuit, puis les épisomes ont été transférés dans la souche P90CN, délétée pour la région lac pro et résistante à l'acide nalidixique. Pour le croisement 0.5 ml d'une culture (en

phase exponentielle) du diploïde approprié et 0.5 ml d'une culture saturée de la souche réceptrice ont été mélangés et incubés au bain-marie (37°C) pendant 15 minutes, puis 30 minutes avec faible agitation. Après titration des recombinants pro⁺nal^R, le mélange a été cultivé dans 2 ml de milieu 2YT pendant 2 heures (atteignant une densité d'environ $1-3 \times 10^9$ cellules/ml) et 0.1 ml d'une dilution 1:10 ont été étalés sur milieu minimal avec Pgal et acide nalidixique, afin de sélectionner les recombinants P⁺. Etant donné que des mutants spontanés I⁻P⁻ peuvent également croître sur ce milieu, 25 à 50 colonies de chaque croisement ont été répliquées sur milieu indicateur McConkey lac. Celui-ci permet de distinguer les niveaux de β -galactosidase produits dans une souche P⁺ (coloration rouge foncé) et P⁻ (rouge clair).

f) Homogénéotisation et guérison de l'épisome (Miller, 1972)

Pour le transfert de certaines mutations ponctuelles sur le chromosome, les épisomes appropriés ont été croisés dans la souche X7800. Quelques colonies indépendantes de ces diploïdes ont été striées sur milieu indicateur Xgal pour détecter les homogénètes I⁻/I⁻ bleus parmi les hétérogénètes blancs. Pour guérir ces homogénètes de leurs épisomes, environ 500 cellules (0.1 ml d'une dilution 10^{-6} de culture saturée) ont été cultivées pendant 24 heures dans 5 ml de LB, pH 7.6, contenant 300 μ g/ml d'acridine-orange. La perte de l'épisome a été déterminée en testant des colonies isolées pour le marqueur pro.

5) Analyse des délétions ton B qui se terminent avant le gène lac I.

a) Construction de dérivés trp R⁻ de souches à délétions ton B - lac I (Miller, 1972)

Grâce au caractère ara⁻ des souches de délétions, la mutation trp R⁻ a pu être introduite facilement par croisement avec un Hfr Cavalli (CSH61) qui porte les marqueurs rapprochés ara⁺ et trp R⁻.

(La contamination de l'arabinose commerciale par du galactose

a nécessité la sélection préalable de révertants de la mutation gal E dans les souches de délétions, comme décrit dans la méthode de localisation à haute résolution.) Les croisements avec la souche Hfr ont été effectués en mélangeant quelques gouttes de cultures exponentielles des souches appropriées, et en les incubant à 37°C pendant 70 minutes. Puis des aliquots de 0.1 ml de différentes dilutions ont été répartis avec 2.5 ml de F-top agar fondu sur milieu minimal avec arabinose, proline, tryptophane et streptomycine. Les dérivés trp R⁻ de plusieurs recombinants ara⁺ ont été déterminés par leur résistance au 5-méthyltryptophane. En effet seules les cellules trp R⁻ avec un niveau constitutif d'enzymes trp peuvent survivre sur un milieu contenant 100 µg/ml de ce composé. Etant donné que la présence de tryptophane dans le milieu interférerait avec l'effet du 5-méthyl-tryptophane, un episome F'trp (CSH33) a été introduit (par répllication) dans les souches de délétions au caractère trp⁻.

b) Dosage de β-galactosidase

Les niveaux de β-galactosidase ont été déterminés par la méthode de Pardee et al. (59) décrite chez Miller (1972). Les souches appropriées ont été cultivées dans du milieu minimal A sans et avec IPTG 10⁻³M. Des aliquots (variant entre 0.05 et 0.5 ml) de ces cultures ont été complétés à 1 ml avec du tampon Z. Après rupture partielle des cellules par le toluène, les suspensions ont été placées dans un bain-marie à 28°C, et additionnées de 0.2 ml d'une solution aqueuse d'o-nitrophenyl-β-D-galactoside (ONPG) à 4 mg/ml. Cette substance étant un substrat de la β-galactosidase, la mesure spectrométrique de la quantité d'o-nitrophénol libéré (par sa coloration jaune) permet de déterminer le niveau de β-galactosidase présent dans les cellules.

c) Préparation d'extraits de cellules (Schmitz et al. 1976)

Les souches appropriées ont été cultivées jusqu'à saturation dans 200 ml 2YT à 37°. Après centrifugation à 5000 x g pendant 10 min. à 4°C, le culot a été resuspendu dans du

tampon TMS à raison de 0.5 ml par 1 mg de cellules, avec 20 μ l d'une solution d'ADNase pancréatique (1 mg/ml, Merck). La suspension a été homogénéisée sur Vortex, puis traitée pendant 20 secondes par un sonicateur Branson B12 pour rompre les cellules. Les débris cellulaires ont été séparés par centrifugation à 30.000 x g pendant 20 minutes à 4°. Le surnageant a été utilisé pour la dialyse à l'équilibre.

d) Dosage du répresseur par liaison avec ^{14}C -IPTG (Miller, 1972)

Les extraits bruts ont été dialysés contre une solution de IPTG (10^{-7}M) marqué au ^{14}C (environ 1000 cpm/0.1 ml, Becton) jusqu'à équilibre pendant au moins 4 heures à 4°C.

La concentration des protéines totales a été déterminée par le réactif de Biuret.

L'activité spécifique du répresseur est définie par le rapport suivant

$$\frac{\% \text{ d'excès de } ^{14}\text{C-IPTG}}{\text{mg de protéines/ml}}$$

BUT DES RECHERCHES PRESENTEES DANS LA THESE

Le travail de cette thèse s'inscrit dans le projet de recherche général de notre laboratoire qui a pour but d'élucider la relation entre la structure et les fonctions du répresseur lac. Notre approche consiste à créer et à analyser un très grand nombre de molécules de répresseur modifiées par la substitution d'acides aminés à différentes positions connues. La majorité de ces modifications ont été obtenues au moyen de la suppression de mutations nonsense dont plus de 5000 ont été isolées pour cette étude (Miller et al. 1977, Coulondre et Miller, 1977a).

L'objectif principal de mon travail a été la construction d'une carte détaillée du gène lacI qui permet de localiser rapidement ces mutations. La première partie de la thèse décrit l'obtention de plus de 400 délétions et leur localisation par rapport aux mutations nonsense et à quelques 800 mutations non supprimibles. Cette analyse a abouti à une carte de plus de 100 intervalles de délétions d'une longueur moyenne de dix paires de bases.

En utilisant cette carte et divers autres moyens génétiques (voir introduction et Miller et al. 1977), un système a été mis au point pour attribuer les sites nonsense à des acides aminés spécifiques du répresseur. La connaissance de la position de ces mutations a alors permis de définir la longueur physique des intervalles de délétions et par conséquent d'établir une corrélation précise entre la carte génétique et la carte physique.

Le haut degré de résolution et la précision de cette carte de délétions ont été particulièrement utiles pour diverses études présentées dans la seconde partie du travail. Elles concernent notamment la distribution des terminaisons de délétions et celle des mutations missense dans le gène lacI, ainsi que les "effets de marqueur".

B) CONSTRUCTION DE LA CARTE DE DELETIONS

1) Sélection de délétions tonB-lacI

L'obtention de délétions dans l'opéron lactose a été hautement facilitée par la construction de souches d'E. coli dans lesquelles l'opéron lactose a été transposé sur un phage transducteur $\phi 80$ dlac au site att80 du chromosome à proximité de l'opéron tryptophane et du site tonB. En effet, une partie importante de mutations dans le site tonB, qui contrôle la sensibilité aux phages T1 et $\phi 80$ et aux colicines V et B, sont des délétions qui s'étendent de part et d'autre de ce site. (Franklin et al. 1965, Yanofsky, Lennox, 1959)

Pour la sélection de délétions tonB-lacI, j'ai utilisé le dérivé (X7800) d'une souche construite à l'origine par Beckwith et Signer (1966). L'orientation de l'opéron lactose dans cette souche est telle qu'il est transcrit dans la même direction que l'opéron tryptophane et que le gène lacI est le plus proche du site tonB. (voir figure 1) Le procédé de sélection de mutants tonB a consisté à traiter des bactéries simultanément par des lysats de $\phi 80$ vir et de colicines V et B. (voir Méthodes) Bien que d'autres sites contrôlent également la sensibilité soit à $\phi 80$, soit aux colicines, seule une mutation tonB confère simultanément à la bactérie la résistance aux 3 facteurs. L'étalement du mélange d'infection sur des boîtes indicatrices avec Xgal a permis de discerner parmi les survivants les souches portant des délétions qui inactivent le gène lacI. De telles cellules produisent un niveau constitutif de β -galactosidase et forment des colonies bleues, alors que les autres survivants restent blancs ou bleu pâle. Aux conditions optimales de sélection, les survivants tonB apparaissent à des fréquences entre 5×10^{-8} et 1×10^{-7} , et en général un sur 500 à 1000 mutants tonB est également I⁻.

Grâce à ce procédé, 408 délétions tonB-lacI spontanées ont été isolées à partir de 600 cultures indépendantes. 161 délétions ont été obtenues par Brigitte Friedli (les numéros

de ces délétions sont précédés par la lettre B) et 247 par moi-même.

2) Distinction des différentes classes de délétions

L'inactivation totale ou partielle du gène lacI peut être causée par des délétions qui se terminent soit dans sa région de contrôle, soit à l'intérieur ou au-delà du gène. (figure 1)

En première approximation, on distingue les deux derniers groupes sur des boîtes indicatrices contenant du Xgal, (avec et sans IPTG).

- I Les délétions qui se terminent à l'intérieur du gène I provoquent une expression entièrement constitutive de l'opéron lactose, ce qui se traduit par la coloration bleu foncé des colonies. 396 délétions de la collection de 408 appartiennent à cette catégorie.
- II Lorsque les délétions éliminent totalement le gène I et endommagent partiellement le promoteur lac, la synthèse de β -galactosidase est constitutive, mais son niveau est réduit. Trois des souches analysées sont de ce type et se distinguent par une coloration bleu clair indépendamment de la présence d'IPTG.

La localisation a permis de vérifier que ces délétions couvrent une mutation ponctuelle connue dans le promoteur lac (voir regroupement des délétions). Leur analyse n'a cependant pas été approfondie car des délétions similaires ont été décrites en détail. (Reznikoff et al. 1969, Mitchell et al. 1975)

- III Pour ce qui est des délétions affectant la région de contrôle du gène lacI, on peut prédire différents phénotypes:

- i) Un dommage important de la région nécessaire à l'initiation de la transcription abolirait l'expression du

gène I et conduirait à une synthèse entièrement constitutive de β -galactosidase.

- ii) Si une telle délétion s'étendait jusqu'à l'opéron tryptophane de façon à placer l'expression du gène I sous contrôle trp, (fusion trp-lacI), un niveau résiduel de répresseur lac pourrait être produit.
- iii) Dans le cas d'une telle fusion, une réduction de la synthèse de répresseur lac pourrait également provenir d'une lésion partielle de la région de contrôle du gène I nécessaire pour la traduction.

Parmi les délétions isolées lors de ce travail, 9 semblaient ne pas affecter le gène de structure du répresseur lac. (Cette constatation repose sur les résultats de recombinaison positifs avec une mutation du quatrième nucléotide du gène I. Ces données seront présentées plus en détail ultérieurement.) Certaines d'entre elles répriment partiellement la synthèse de β -galactosidase, les autres expriment cette enzyme constitutivement.

Afin de mieux définir cette nouvelle catégorie de délétions, j'ai testé l'influence d'une dérégulation de l'opéron tryptophane sur l'expression du gène I dans ces souches.

a) Caractérisation des délétions affectant la région de contrôle du gène I

Une mutation trpR⁻, qui provoque l'expression constitutive de l'opéron tryptophane, a été introduite dans les 9 souches de délétions tonB-lacI décrites plus haut. Pour évaluer l'effet de cette mutation sur la synthèse du répresseur lac, j'ai mesuré le taux de β -galactosidase ainsi que l'activité de répresseur (par liaison avec l'IPTG marqué au ¹⁴C) pour chacune des 9 délétions dans des souches trpR⁺ et trpR⁻. Les résultats de ces analyses, présentés au tableau 1, révèlent que dans les souches portant les délétions 101, 119, 43 et 27, la mutation trpR⁻ augmente de façon significative le taux de répression de l'opéron lactose dans le cas

de la délétion B80, l'effet mesuré est plus faible. Ces résultats sont confirmés pour les délétions 101 et 43 par l'activité nettement plus élevée du répresseur lac en présence de la mutation trpR⁻. Dans le cas des délétions 119, 27 et B80, les activités de répresseur sont si proches de la limite de détection qu'il est difficile de voir une différence entre l'état trpR⁺ et trpR⁻. Toutefois, il faut mentionner que la quantité de répresseur produit par un gène I⁺ n'est pas détectable par la technique de dialyse utilisée dans ces expériences.

Les résultats obtenus avec la délétion B116 étaient très surprenants. En particulier le niveau de répresseur (déterminé par liaison avec IPTG) dans une souche trpR⁻ atteint une valeur environ 150 fois plus élevée que dans une souche type sauvage. La constante de liaison pour l'IPTG du produit de B116 étant identique à celle du répresseur normal, la quantité maximale de répresseur B116 représente donc de 0.5 à 1% des protéines solubles. La cause de cette production exceptionnelle a été élucidée grâce à l'analyse de la séquence des acides aminés du produit de B116 par H. Sommer et H.G. Wittmann. Ces résultats sont décrits dans l'annexe. Quant à l'activité partiellement constitutive de β -galactosidase, dans la souche trpR⁻, elle semble être le résultat d'une transcription plus efficace au-delà du gène lacI jusque dans l'opéron lac (détails voir annexe).

Il apparaît donc que dans 6 des 9 souches testées le contrôle de l'opéron tryptophane s'exerce, à des degrés divers, sur la production du répresseur lactose. Ceci suggère que les délétions en question éliminent totalement ou partiellement la région de contrôle du gène lacI d'un côté et au moins le site de terminaison pour la transcription de l'ARN messager polycistronique trp de l'autre. De cette façon la transcription commencée dans l'opéron tryptophane peut continuer directement dans le gène lacI. Les délétions 101, 119 et 43 se terminent vraisemblablement à l'intérieur des gènes de structure de l'opéron trp car les souches respectives sont incapables de produire du tryptophane. Par contre, les délétions

27 et B80 semblent laisser l'opéron trp intact. Comme il a été suggéré par les travaux de Mitchell (1975, 1976) que le site de terminaison pour la transcription de l'opéron trp se trouve au-delà du dernier gène de structure (trpA), il est raisonnable de penser que les délétions 27 et B80 s'arrêtent dans la région postcistronique juste avant le terminateur.

Des fusions de ce type peuvent également s'exprimer au niveau de la traduction. En effet il est possible de créer des protéines chimériques contenant tout ou en partie du produit d'un gène de l'opéron trp et du répresseur lac, à condition que la lecture puisse continuer en phase d'un élément à l'autre et qu'elle ne soit pas interrompue par un codon non-sense. Jusqu'à présent seule la séquence amino-terminale du produit de la délétion B116 a été déterminée. (Sommer et al. 1978) Les propriétés exceptionnelles de cette délétion et du répresseur chimérique résultant seront brièvement décrites dans l'annexe.

b) Extrémités des délétions hors du gène lacI

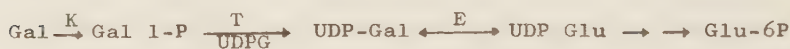
L'emplacement du début des délétions tonB-lacI n'a pas été étudié avec précision. Néanmoins, comme l'opéron tryptophane se trouve à proximité de tonB, il a paru intéressant de déterminer quelle fraction des délétions tonB-lacI affecte les gènes de structure trp. 59% des souches testées sont incapables de croître sur milieu minimal sans tryptophane. Les délétions correspondantes doivent donc inactiver au moins le dernier gène de l'opéron, trpA. Le tableau 2 présente le caractère trp pour toutes les délétions. Ces dernières sont indiquées par ordre numérique pour faciliter la lecture. L'analyse dans l'ordre des groupes de délétion n'a d'ailleurs révélé aucune relation évidente entre le caractère trp et la position des terminaisons de délétions dans le gène lacI.

3) Localisation des délétions

Le procédé général de localisation a consisté à introduire les diverses mutations I⁻ situées sur l'épisome F'lacpro dans les souches F⁻ portant les délétions tonB-lacI et à tester

ces diploïdes pour leur capacité de produire des recombinants I⁺.

Pour la sélection des recombinants I⁺, j'ai utilisé des adaptations de la méthode décrite par Davies et Jacob (1968). Celle-ci repose sur l'observation que des cellules déficientes en galactose-épimérase (galE) lysent en présence de galactose. La cause exacte de cet effet n'est pas élucidée. Il semblerait que l'accumulation d'UDPgal (voir métabolisme du galactose ci-dessous) inhibe la formation de précurseurs essentiels d'une couche de la paroi cellulaire. (Fukasawa, Nikaido, 1961)



Métabolisme du galactose (d'après Yarmolinsky et al. 1961)

Par analogie, le galactose formé par hydrolyse d'un substrat approprié de la β -galactosidase peut tuer des souches portant une mutation galE. Si l'on utilise un substrat non-inducteur tel que le phényl- β -D-galactoside à une concentration de 0,5 mg/ml, les cellules contenant un niveau constitutif de β -galactosidase meurent, tandis que les souches I⁺ survivent. Toutefois deux types de mutants spontanés résistent aussi à cette sélection, notamment les "révertants" galK et galT, ainsi que les mutants lacZ. Les techniques de localisation décrites dans la section suivante permettent d'éviter que ces mutants indésirés n'interfèrent avec la détermination des recombinants I⁺.

a) Méthode de localisation à haute résolution

Comme les 408 délétions tonB-lacI ont été isolées dans une souche portant une mutation galE, la sélection par Pgal a pu être appliquée directement aux diploïdes F[']I⁻/Δ(tonB-lacI). Pour chaque cas, 4 à 5x10⁷ cellules ont été testées sur des boîtes à milieu minimal avec glucose, tryptophane et Pgal.

Les 3 types de survivants apparaissant sur ce milieu ont pu être différenciés par répllication sur milieu minimal glucose- trp- Xgal:

- Etant donné que les mutants galK et galT ont maintenu un niveau constitutif de β -galactosidase (lacI⁻), ces colonies se distinguent par leur coloration bleu foncé. En général, de tels mutants apparaissent à raison de 50 à 100 colonies par boîte, c'est-à-dire à une fréquence de 1 à 2×10^{-6} .
- Les colonies avec des mutations Z⁻ qui inactivent totalement la β -galactosidase sont entièrement blanches. Celles-ci étaient cependant très rares (de l'ordre de 10^{-8}) car les diploïdes testés contenaient deux copies du gène Z.
- Les recombinants I⁺ représentent la troisième catégorie de survivants. Ils forment des colonies bleu pâle, grâce à leur niveau résiduel de β -galactosidase.

Afin de distinguer ces derniers définitivement des mutants Z⁻, le caractère inductible de quelques colonies I⁺ a été vérifié dans chaque cas par répllication sur milieu indicateur Xgal sans et avec IPTG (10^{-3} M).

Au cours de la localisation, des recombinants I⁺ ont été observés à des fréquences allant de 10^{-7} jusqu'à 10^{-3} suivant les délétions et mutations employées.

b) Limites de résolution

La possibilité de détecter des recombinants à des fréquences de 10^{-8} à 10^{-7} par la méthode décrite ci-dessus a permis dans plusieurs cas de séparer des mutations I⁻ distantes de quelques bases seulement. Le critère employé pour une telle séparation repose sur la capacité d'une délétion de former des recombinants I⁺ à une fréquence détectable avec un seul des deux marqueurs. Une fréquence inférieure à 10^{-7} a généralement été considérée comme un résultat négatif. Quelques cas typiques sont représentés dans la figure 2 et le tableau 3. Par convention, les terminaisons de délétions ont été placées entre les sites de mutations. Il est cependant possible qu'une homologie de quelques bases entre une délétion et une mutation ponctuelle soit indispensable pour qu'une recombi-

naison ait lieu, et dans ce cas, les délétions s'arrêteraient effectivement avant les deux mutations très rapprochées. Ceci ne devrait pourtant avoir aucune influence sur l'ordre des mutations. Les fréquences exactes de recombinants obtenus par croisement des différentes mutations ponctuelles et délétions de la figure 2 sont indiquées au tableau 3.

Certains résultats de très basses fréquences de colonies I^+ ont été vérifiés par l'analyse d'un plus grand nombre de cellules. En effet seule une différence significative par rapport aux colonies I^+ dues à la réversion permet de définir s'il s'agit de recombinants. Les fréquences de réversion ont été déterminées dans des diploïdes avec des délétions qui couvraient les mutations ponctuelles en question. Dans chaque cas, environ 10^9 cellules ont été analysées sur milieu sélectif. (voir méthodes) Le résultat est considéré comme positif si la fréquence de recombinants est 10 fois plus élevée que la fréquence de révertants. (voir tableau 4) Quelques cas spéciaux, pour lesquels la fréquence des recombinants se situe très près de la limite de détection, seront traités à part (chapitre sur le regroupement des délétions)

c) Méthode de localisation par réplication

Pour un premier regroupement des 408 délétions, j'ai bénéficié de 52 mutations ponctuelles I^- dont l'ordre avait pour la plupart, été établi préalablement par J. Miller et D. Ganem. Réalisé par la méthode à haute résolution, cette localisation aurait nécessité près de 20.000 croisements individuels. Par contre, la méthode de réplication élaborée par J. Miller m'a permis de placer chaque délétion dans un des intervalles définis par ces mutations en analysant les 52 diploïdes simultanément sur une boîte de Petri.

Etant donné que dans le procédé par réplication une mise en évidence des différents types de survivants de la sélection galE n'est pas possible, la présence d'autres colonies que les recombinants I^+ mènerait à des résultats inexacts.

Pour éliminer les mutants I^-Z^- , I^-galK^- ou I^-galT^- , l'o-nitrophényl- β -D-thiogalactoside (TONPG) a été choisi. Ce composé n'induit pas l'opéron lactose, mais lorsqu'il est concentré dans des bactéries par un haut niveau constitutif de lac perméase, (produit du gène lacY), il inhibe la croissance de ces cellules sur du succinate, par un mécanisme inconnu. (Smith et Sadler, 1971) Cette sélection est cependant imparfaite dans des souches tonB, et est complétée par la sélection sur milieu Pgal-glucose. Cette dernière doit être effectuée séparément car le glucose interfère avec l'action de TONPG.

Pratiquement la technique (décrite sous Méthodes) consiste à effectuer environ 50 croisements simultanément en répliquant les souches $F'I^-$ sur un tapis de $F^-\Delta tonB-I$. (le procédé inverse, soit la répllication des souches $F^-\Delta tonB-I$ sur les souches $F'I^-$ peut également être appliqué). Ensuite les diploïdes sont répliqués successivement sur des boîtes à milieu minimal succinate+trp +TONPG et sur milieu minimal glucose-trp-Pgal.

d) Résolution de la méthode de répllication

Quelques résultats typiques obtenus par cette méthode sont illustrés par 6 exemples dans la figure 3. Une série de 52 mutants disposés dans l'ordre correspondant aux positions des mutations dans le gène lacI a été croisée avec 6 délétions différentes. Sur les boîtes Pgal-glucose, représentées dans les photographies, les taches noires représentent les recombinants I^+ . Pour chaque point décisif, les résidus affectés par les mutations I^- correspondantes ont été indiqués sur la figure 4. Si l'on compare les figures 3 et 4, il apparaît clairement que le procédé de répllication permet de différencier facilement des mutations séparées par 10 nucléotides et plus. Dans un cas particulier (figure 3a) deux marqueurs affectant des acides aminés adjacents (en position 6 et 7), à une distance de 4 nucléotides, ont pu être distingués. La délétion 208 recombine en fait avec la mutation

correspondant au résidu 6 à une fréquence très basse (de 2.5×10^{-7} , voir tableau 3), mais celle-ci n'est pas détectée par le procédé de répliation. Une détermination par la méthode à haute résolution a montré que la fréquence de recombinants entre la délétion 208 et le marqueur du résidu 7 s'élève à 10^{-6} . Ceci doit donc être considéré comme la limite de résolution de la méthode de répliation.

e) Regroupement des délétions

La carte de délétions détaillée finale, comprenant plus de 100 intervalles distincts, a été construite en trois étapes principales:

I Tout d'abord les 408 délétions ont été mappées par la méthode de répliation contre 52 mutations ponctuelles I⁻. Parmi celles-ci figuraient 48 mutations nonsense, isolées par J. Miller et C. Coulondre à partir de cultures traitées avec différents mutagènes (mutT, 2AP, UV). Leurs positions respectives avaient été déterminées préalablement grâce à la série de délétions X8600 (Miller et al. 1970, Ganem, 1972) et à des analyses de recombinaison entre mutations ponctuelles (Coulondre et Miller 1977a). Cette série comprenait également quatre mutations missense dont les résidus affectés dans la protéine (à savoir 5, 16, 19 et 53) avaient été élucidés par analyse peptidique. (Weber, 1975)

Les résultats obtenus par répliation ont été vérifiés par la méthode à haute résolution en croisant chacune des délétions avec les deux mutations ponctuelles délimitant l'intervalle correspondant. De cette manière les 408 délétions ont été réparties en 41 groupes distincts, comprenant entre 1 et 47 délétions (figure 5a-c).

II L'étape suivante consistait à tester toutes les délétions de chaque groupe avec de nouvelles mutations nonsense, induites en grande partie par des mutagènes tels que NG, EMS et NQO. Ces mutations avaient été différenciées de celles utilisées préalablement par divers autres critères (par exemple la spécificité des mutagènes, l'efficacité

de suppression, voir Miller et al. 1977, Coulondre et Miller, 1977a). Pour ce faire les souches portant les délétions ont été répliquées sur des boîtes couvertes d'un tapis de la souche F'I⁻ adéquate, puis les recombinants I⁺ ont été sélectionnés de la façon habituelle. Cette analyse a été effectuée avec 32 mutations nonsense supplémentaires et chaque nouvel intervalle a été vérifié par la localisation à haute résolution. Ainsi le nombre de groupes de délétions a été porté à 64 (figure 6 a-d).

C'est sur la base de cette localisation et grâce au système de corrélation élaboré par J. Miller que la majorité des mutations nonsense indiquées dans la figure 6 ont pu être attribuées à un résidu précis du répresseur lactose. (Miller et al. 1977, Coulondre et Miller, 1977a). La liste des marqueurs utilisés et des acides aminés correspondants figure au tableau 5. Les positions exactes de plusieurs de ces mutants ont par ailleurs été déterminées par analyse peptidique ou analyse de séquence d'ADN (détails voir légende de la figure 6).

La connaissance de ces positions a permis d'établir une corrélation entre la carte génétique et la carte physique du répresseur lactose qui sera présentée plus loin.

III La limitation des sites nonsense aux codons dont une seule base (ou dans de rares cas 2 bases) diffère des triplets UAG, UAA ou UGA m'a conduit à compléter la carte avec une collection de plus de 600 mutations missense isolées par J. Miller. Cette série comportait des mutations spontanées et induites par les mutagènes 2AP, NQO, UV et mutT. Leur localisation précise qui sera présentée en détail ultérieurement a permis de définir 35 nouveaux intervalles.

Diverses autres subdivisions ont été effectuées pour aboutir à la carte de délétions définitive représentée à la figure 7 (a-b). Etant donné qu'aucune mutation nonsense n'a pu être isolée dans la partie du gène I correspondant aux 35 résidus de la partie C-terminale du répresseur, le dernier

intervalle a été subdivisé par d'autres moyens:

Les délétions ont été séparées en deux groupes (99 et 100) par croisement avec la délétion L1, qui coupe les 27 derniers nucléotides du gène I. (Gilbert et al. non publié) Puis toutes les délétions allant au-delà de la terminaison de L1 ont été croisées (voir Méthodes) avec la mutation du promoteur lacL8, qui est située à 19 bases du codon de terminaison du répresseur (Maxam, Sundaresan, non publié). Avec trois de ces délétions, il n'a pas été possible de reconstituer un promoteur lac sauvage par recombinaison. De plus ces délétions diminuent la synthèse maximale de β -galactosidase alors que les autres délétions se terminant avant L8 ne semblent pas affecter l'expression de l'opéron lac.

Quant à la subdivision du groupe 99 entre le dernier marqueur nonsense et L1, elle est fondée sur les différences significatives des fréquences de recombinants observées dans cet intervalle. Etant donné que dans ces cas il ne s'agit pas d'une séparation proprement dite, les délétions ont été réparties dans des sous-groupes désignés par A, B, C ... Des valeurs typiques pour chacun des sous-groupes ainsi définis figurent au tableau 6. Des résultats similaires ont conduit également à la subdivision du groupe de délétions 2.

L'attribution de ces délétions à différents groupes semble justifiée par l'observation qu'il existe une relation linéaire entre la fréquence de recombinants et la distance séparant une mutation ponctuelle de diverses délétions. (Cette étude sera présentée en détail ultérieurement).

Par ailleurs des tests de réversion ont été réalisés pour tous les croisements où la fréquence de recombinants est très basse. Les résultats cités au tableau 7 ont permis une séparation en deux intervalles de délétions distincts. Dans deux autres cas, la différence entre recombinants et révertants semblait également significative, mais les fré-

quences de recombinauts se situaient si près de la limite de résolution que les délétions correspondantes ont été réparties dans des sous-groupes, notamment 29 A, 29 B et 96 A, 96 B. La subdivision du groupe 38 est un cas particulier, dû à l'influence de différentes substitutions de bases au même site sur la fréquence de recombinaison. (Ce phénomène sera traité spécialement dans l'étude sur les effets de marqueurs) Lorsque la délétion 190 est mappée contre le dérivé amber (A14) de la tyrosine en position 126, des recombinauts sont détectés à une faible fréquence, mais avec le dérivé ochre (O18) du même codon le résultat est négatif. (tableau 7)

Pour la carte détaillée, j'ai utilisé en outre la délétion 626 de la série X8600 de J. Miller, qui sépare les marqueurs A19 et O21, car une telle délétion ne figurait pas dans ma collection.

f) Carte de délétions détaillée

Toutes les analyses décrites ont finalement abouti à une division des 408 délétions en 108 groupes, dont 106 se trouvent à l'intérieur du gène lac I (figure 7 a-b). Les numéros 1 à 100 ont été attribués aux intervalles définis clairement par une mutation nonsense ou missense. Certains intervalles comprennent des sous-groupes, désignés par A, B etc, dont la détermination a été décrite.

Chacune des mutations indiquée sur la figure 7 a été localisée par la méthode à haute résolution. Celles inscrites dans les cadres ont été croisées avec toutes les délétions des deux groupes adjacents. Les autres mutations ont été testées seulement avec la première délétion des groupes voisins. Les mutations ponctuelles indiquées dans cette carte ne représentent qu'une partie de toutes les mutations employées pour sa construction. L'ensemble des mutations nonsense a été décrit par Miller et al. (1977; Coulondre et Miller, 1977a). La distribution des quelques 600 mutations missense sera présentée dans une section ultérieure.

4) Corrélation de la carte génétique avec la carte physique

106 intervalles distincts dans un gène de 1080 nucléotides représentent sans doute une forte capacité de résolution. Toutefois l'innovation et l'avantage majeurs de ce système résident dans la corrélation de la carte de délétions avec la carte physique du gène lacI.

66 des 106 intervalles du gène sont délimités par des mutations nonsense (voir figure 6) et pour la majorité de celles-ci, l'acide aminé correspondant dans le répresseur a été déterminé. (Miller et al. 1977, Coulondre et Miller 1977a) A ces marqueurs s'ajoutent les trois mutations missense déterminées par analyse peptidique. Ainsi, ces groupes de délétions ont pu être corrélés avec des intervalles précis dans la carte physique du répresseur, ce qui est représenté dans la figure 8 (a-b). Il ressort clairement de cette carte que la corrélation des mutations nonsense permet également une bonne estimation des positions de nombreuses mutations missense. Toute nouvelle mutation du gène lacI peut donc être attribuée rapidement à une région assez restreinte du répresseur.

De plus, cette corrélation entre mutation nonsense et résidus du répresseur a pu être étendue à la position des paires de bases affectées dans le gène lacI. La détermination de ces positions par des moyens génétiques et grâce à l'élucidation de la séquence d'ADN du gène lacI, (Farabaugh, 1978) a finalement permis d'exprimer les longueurs des intervalles en nombre de nucléotides. Cette précision a été particulièrement utile pour analyser la distribution des terminaisons de délétions à l'intérieur du gène I, qui sera décrite ci-dessous.

5) Distribution des terminaisons de délétions dans le gène lacI

Les études sur la distribution des terminaisons de délétions dans les opérons gal et ara chez E. coli ont révélé

deux classes distinctes déterminées par le mode de sélection. Des sites de terminaison préférentiels ont été trouvés dans l'opéron gal pour des délétions formées par excision aberrante d'un prophage λ (Pfeifer, Rosypal, 1969, Pfeifer et al. 1974) ou d'une séquence IS1 (Reif et Saedler, 1975, 1977). Par contre une étude génétique de l'opéron ara a suggéré que des délétions spontanées s'y terminent au hasard. (Schleif, 1972)

Dans le gène lacI une telle analyse de distribution était particulièrement propice grâce à la multitude de marqueurs connus qui permettaient d'examiner les terminaisons des délétions dans des intervalles d'environ 15 nucléotides. L'étude a porté sur 396 délétions dont les terminaisons sont comprises dans un segment de 1098 nucléotides, entre la mutation O1 (affectant le 4^e nucléotide du gène I, Coulondre et Miller, 1977a) et L8 (situé à 19 nucléotides après le codon de terminaison pour le répresseur, Maxam et Sundaresan, non publié). Ce segment a été divisé en 51 intervalles par des marqueurs connus, puis le nombre de terminaisons de délétions dans chaque groupe a été normalisé à des intervalles réguliers de 15 nucléotides (tableau 8). La valeur moyenne normalisée de délétions par intervalle était de 6.

Afin de déterminer dans quelle mesure la distribution expérimentale était due à l'effet du hasard, une distribution de Poisson a été calculée par la fonction suivante:

$$P = \frac{e^{-a} a^n}{n!}$$

où a représente la valeur moyenne et n le nombre de terminaisons de délétions par intervalle.

Cette fonction étant telle que $\sum_{n=0}^{\infty} P=1$, la distribution expérimentale a été représentée de façon à ce que la somme des intervalles considérés soit égale à 1. Dans la figure 9 les distributions observées (histogramme) et calculées (ligne continue) ont été superposées. Il apparaît clairement que

la plupart des délétions se terminent au hasard à l'intérieur du gène lacI.

Néanmoins dans deux intervalles le nombre de délétions semble dépasser largement la valeur théorique. Ces deux groupes ont été éliminés et une distribution de Poisson a été calculée pour la nouvelle moyenne de 5.5 (ligne pointillée, figure 9). Bien que dans ce cas deux autres intervalles semblent se trouver à la limite d'une distribution statistique, la conformité entre ces distributions théorique et expérimentale a été déterminée selon le test de χ^2 . La valeur calculée de $\chi^2 = 5.919$ est légèrement inférieure à la valeur théorique (pour deux degrés de liberté et une limite de tolérance de 5%), soit $\chi^2_{P=0.05} = 5.991$. Par convention ceci indique que les écarts trouvés sont dus à la variabilité statistique. Dans l'ensemble les terminaisons dans les 49 intervalles considérés sont donc distribués statistiquement.

Toutefois, si l'on veut évaluer l'importance des intervalles "chauds" individuels, il faut tenir compte de l'influence de la normalisation sur ces résultats. Il s'agit alors de déterminer quelles seraient les probabilités dans chaque cas si des intervalles d'une longueur différente avaient été utilisés. Pour chacun des intervalles situés hors des limites (figure 9) une distribution de Poisson a été calculée avec les nombres non normalisés. Au tableau 8 bis figurent les valeurs respectives de $(1 - \sum_{n=0}^{k-1} P)$, qui correspondent aux probabilités de trouver un intervalle avec un nombre de terminaisons de délétions égal ou supérieur au nombre k en question. Ces valeurs indiquent qu'un groupe tel que I est à la limite de la distribution statistique, alors que les trois autres régions peuvent être considérées comme particulièrement propices à la formation de délétions. Malheureusement l'autre extrémité des délétions de ces trois groupes n'est pas connue et ceci limite sérieusement toute conclusion à propos de leur origine.

C) APPLICATIONS DE LA CARTE DE DELETIONS

1. Fréquences de recombinaison en fonction de la distance entre délétions et mutations ponctuelles

Les études publiées sur la relation entre la fréquence de recombinants entre 2 mutations et la distance qui les sépare n'ont été effectuées que pour des paires de mutations ponctuelles (Yanofsky et al. 1964, Celis et al. 1973). Aussi était-il intéressant d'entreprendre une étude similaire de recombinaison soit entre une délétion et des mutations localisées et situées à des distances croissantes, soit entre une mutation et plusieurs délétions dont les terminaisons sont de plus en plus éloignées.

Pour la première étude, j'ai choisi la délétion B18 (qui semble se terminer peu avant la mutation affectant le résidu 5 et une série de 20 mutations nonsense situées à des distances allant de quelques nucléotides jusqu'à environ 850. Pour certaines de ces mutations ponctuelles, la position exacte dans la protéine correspondante a été déterminée par analyse de la séquence du répresseur altéré. Il s'agit des résidus 5, 7, 16, 26, 60, 201 et 220. (Weber et al. 1975, Sommer, 1977) Pour les autres mutations, la corrélation a été établie au moyen de la méthode génétique décrite par Miller et al. (1977) et Coulondre et Miller (1977a) (voir tableau 5).

3 à 10 croisements ont été effectués pour chaque mutation et les résultats combinés sont présentés au tableau 9 et à la figure 10. Sur l'ensemble du graphique on peut discerner une tendance générale de la fréquence de recombinants à augmenter avec la distance, mais de façon très irrégulière. De telles déviations de la linéarité appelées "effets de marqueurs", ont été relevées lors de nombreuses études sur la recombinaison entre mutations ponctuelles: (Yanofsky et al. 1964, Drapeau et al. 1968, Norkin, 1970, Stadler et Kariya, 1973, Moore et Sherman, 1975, 1977). Un effet particulièrement frappant a été observé pour la recombinaison de la dé-

létion avec l'amber et l'ochre dérivés du codon spécifiant la tyrosine en position 126 (figure 10). Ceci m'a conduit à étudier de façon plus systématique l'influence de différentes substitutions dans un même codon sur la fréquence de recombinaison.

2) Etude "d'effets de marqueurs"

Afin de tenter de dégager une relation précise entre les effets de marqueurs et des substitutions de bases spécifiques, j'ai comparé les fréquences de recombinants obtenus pour les croisements d'une délétion avec les mutations amber et ochre dérivés d'un même codon à 15 sites différents. Etant donné que les interférences s'exercent le plus fortement sur la recombinaison entre mutations rapprochées, j'ai utilisé des délétions qui se terminent à proximité des mutations en question.

2 types distincts de paires ochre-amber ont été testés.

a) Mutations ochre et amber dérivées de codons spécifiant Gln, Glu et Leu

Dans un cas, des mutations ochre (UAA) ont été dérivées des codons CAA (glutamine) GAA (acide glutamique) et UUA (leucine), puis converties en amber (UAG) (Coulondre et Miller, 1977a).

En considérant les résultats pour cette série de mutations, (tableau 10), il faut cependant tenir compte du fait que le codon amber (UAG) ainsi créé diffère du triplet original par deux bases. Ainsi les fréquences de recombinants inférieures observées avec la plupart des mutations amber peuvent s'expliquer par un appariement plus difficile et ne permettent donc pas de conclusions quant à un effet spécifique.

b) Mutations ochre et amber dérivées des codons de la tyrosine

Le seul système permettant d'étudier, avec l'approche choisie, l'effet de différentes substitutions de bases à un site précis est représenté par les mutations dérivées des

codons de la tyrosine, à savoir UAU et UAC.

J'ai testé les mutations ochre et amber isolées indépendamment à partir de chacun des 8 codons spécifiant la tyrosine dans le gène I, et dans plusieurs cas, les mutations amber obtenues par conversion des ochres. Les valeurs moyennes de 3 à 9 déterminations pour chaque croisement sont présentées au tableau 11.

Une comparaison des fréquences de recombinants met en évidence deux classes distinctes:

- I. Dans 3 cas, les fréquences sont quasi identiques pour les mutations amber et ochre.
- II. Dans la seconde catégorie, représentée par les 5 autres tyrosines, les mutations amber présentent une fréquence de recombinants 5 à 10 fois plus élevée que la mutation ochre correspondante.

La connaissance des codons spécifiant les diverses tyrosines (grâce au système mis au point par Miller et al. 1977) a permis d'établir une corrélation intéressante entre une substitution de base spécifique et son influence sur la recombinaison: En effet, les membres de la catégorie I sont représentés par le codon UAC et tous ceux de la classe II par UAU.

Il est important de s'assurer que la différence des fréquences de recombinants ne soit par un artéfact dû à une apparition tardive des recombinants dans l'un des diploïdes et une longue accumulation dans l'autre. A cet effet, 3 cultures indépendantes des diploïdes portant respectivement l'amber et l'ochre dérivés de la tyrosine 126 ont été cultivés pendant 10 à 15 générations supplémentaires. Des variations d'un facteur de 2 au maximum ont été observées pour les fréquences de recombinants respectives, suggérant que les différences indiquées au tableau 11 proviennent réellement des effets de marqueur. Par ailleurs, l'examen de la séquence de bases voisines aux codons en question n'a révélé aucune analogie. Il semble donc bien que la substitution spécifique d'une

seule base soit capable d'influencer aussi fortement le phénomène de recombinaison.

Lorsqu'on croise les mutations **amber** et **ochre** (au site de la tyrosine 126) avec différentes délétions qui se terminent à des distances croissantes, on s'aperçoit que l'effet de marqueur s'atténue considérablement en fonction de la distance. Le rapport des fréquences de recombinants **amber/ochre** est d'environ 10 lorsque la distance entre la mutation et la délétion ne représente que quelques bases, mais diminue rapidement pour atteindre une valeur autour de 2 à partir d'une distance de 100 jusqu'à 350 nucléotides. (tableau 12).

La représentation graphique de ces résultats à la figure 11 met d'ailleurs en évidence une dépendance linéaire entre les fréquences de recombinants et la distance séparant une mutation donnée et diverses délétions. Ceci est une bonne indication que les effets de marqueurs sont reliés directement aux mutations ponctuelles. Par conséquent, il est dangereux d'établir un ordre pour différentes mutations ponctuelles à l'intérieur d'un même intervalle de délétions sur la base des fréquences de recombinants. Par contre cette observation m'a permis de subdiviser un certain nombre de délétions d'un même groupe grâce aux différences significatives entre leurs fréquences de recombinaison avec une même mutation (voir Regroupement des délétions).

3. Mutations I^- non suppressibles

Afin d'améliorer la résolution de la carte de délétions créée par les mutations nonsense, près de 800 mutants I^- autres que **amber** ou **ochre** ont été isolés parallèlement à partir de cultures soit non traitées, soit traitées par des mutagènes tels que mutT, 2AP, 4NQO et UV. La localisation de toutes ces mutations, ainsi que diverses analyses complémentaires ont permis de distinguer parmi ces mutations I^- les différentes classes présentées au tableau 13.

a) Distribution des mutations missense

Tous les mutants pour lesquels j'ai pu démontrer un caractère différent d'une mutation ponctuelle I^- (groupes b, c, d, tableau 13) vont être traités séparément plus loin.

La localisation et l'analyse de la distribution des mutations missense porte sur les 606 mutations considérées comme ponctuelles. Il faut toutefois noter que parmi ces dernières peuvent figurer quelques mutations "frameshift", à savoir de petites délétions ou insertions que les méthodes utilisées ne peuvent déceler.

Les mutations I^- ont d'abord été groupées selon la capacité de complémentation négative des répresseurs correspondants, car tous les mutants connus avec des propriétés transdominantes affectent les 60 à 70 premiers acides aminés du répresseur. Pour ce faire, les épisomes porteurs de mutations I^- ont été transférés dans les souches F^- appropriées (portant soit le gène I sauvage soit la mutation I^Q) puis les diploïdes ont été répliqués sur des boîtes indicatrices Xgal. Ceci permet de distinguer les mutants I^{-d} , capables d'inactiver le répresseur lac en trans (colonies bleues) des mutants I^- récessifs (colonies bleu pâle).

Ensuite toutes ces mutations missense ont été croisées avec les 408 délétions par la méthode de réplification. Une grande partie des résultats a été confirmée par le procédé de localisation à haute résolution, en particulier dans tous les cas où une de ces mutations créait un nouvel intervalle de délétions.

L'ensemble des résultats est présenté dans des cartes détaillées figures 12(a-h), 13(a-h), 14(a-h). Les mutations sont indiquées séparément selon leur provenance mutagénique. Les groupes distincts à l'intérieur d'un même intervalle de délétions ont été déterminés principalement par deux moyens:

- Dans la région correspondant aux 60-70 résidus de la partie amino-terminale du répresseur où sont localisés tous les mutants I^{-d} , les différents degrés de complémentation négative que l'on peut discerner sur les boîtes indicatrices Xgal ont été utilisés pour subdivi-

ser certains groupes.

- Dans la seconde partie du gène I, ce sont surtout les résultats de recombinaison entre mutations ponctuelles qui ont permis de définir si les mutations d'un seul groupe affectaient le même site ou non.

Pour 4 points chauds, une mutation missense a été transférée sur le chromosome (par homogénéotisation, voir Méthodes) puis croisée avec tous les autres membres du groupe. Si la fréquence de recombinants était inférieure à 2×10^{-8} , les mutations en question étaient considérées comme identiques. De cette façon, les "points chauds" suivants ont été confirmés: All à A30, les 25 mutations induites par 2AP localisées dans le même intervalle que X7 et les 11 mutations induites par UV situées dans l'intervalle immédiatement après X10 (figures 12, 13).

En ce qui concerne les 18 mutations spontanées situées dans l'intervalle correspondant à la région entre les résidus 193 et 201 de la protéine, elles ont également été attribuées à un seul site. En outre, l'analyse de séquence (Farabaugh, 1978) a révélé qu'il s'agit dans ce cas d'une délétion de 4 bases (CTGG) qui sont répétées 3 fois en tandem dans la séquence originale.

Les subdivisions supplémentaires des groupes de mutations sont fondées sur les résultats de recombinaison (positif ou négatif) avec des mutations nonsens sur le chromosome. (Cette dernière étude visait à détecter des délétions internes du gène lacI.)

La multitude des intervalles de délétions définis par des marqueurs à positions connues m'a permis de délimiter assez étroitement les régions correspondantes du répresseur touchées par les mutations missense. Pour localiser les mutations inconnues qui définissaient un nouvel intervalle, j'ai calculé les distances d'après le nombre de délétions se terminant entre les marqueurs connus et les mutations en question.

Ainsi, les distributions des mutations missense induites par les différents mutagènes ont pu être confrontées avec la carte physique du répresseur, ce qui est représenté à la figure 15.

Une des caractéristiques les plus frappantes de ces distributions est la présence de "points chauds", c'est-à-dire de sites préférentiels pour l'induction de mutations. Ceux-ci sont en général liés à la séquence de l'ADN plutôt qu'à une particularité dans le polypeptide correspondant. Leur apparition à différents endroits selon le mutagène suggère que ces sites favorisent le mécanisme d'action de chacun des mutagènes.

Une telle relation directe entre des "points chauds" de mutations et une particularité de l'ADN a été démontrée récemment pour des mutations amber spontanées et induites par 2AP dans le gène lacI. (Coulondre et al. 1977a) En effet, dans les 3 points chauds observés, il s'agit de la substitution d'une cytosine méthylée comprise dans la séquence ^{CCAGG}GGTCC.

b) Délétions internes du gène lacI

La localisation des mutations I⁻ avec des délétions tonB-lacI qui entrent toutes du même côté dans le gène lacI ne permet pas de détecter d'éventuelles délétions internes. Aussi ai-je croisé les quelques 800 mutations isolées sur l'épisome F'lacpro avec des marqueurs nonsense appropriés, qui avaient été transférés sur le chromosome par homogénéisation (voir Méthodes). (Coulondre et Miller, 1977a) Toutes les mutations non-suppressibles qui donnaient un résultat de recombinaison négatif avec au moins un marqueur par la méthode de réplification ont été testées contre plusieurs marqueurs voisins par le procédé à haute résolution. En utilisant une série de 34 marqueurs, j'ai décelé 37 délétions, dont 15 sont entièrement comprises à l'intérieur du gène lacI. (figure 16)

c) Mutations à haute fréquence de réversion

Pour toutes les mutations qui semblaient donner un résultat

de recombinaison positif avec le dernier groupe de délétions tonB-lacI, la fréquence de réversion a été testée. Dans le cas des mutations citées au tableau 14, des révertants ont été observés à des fréquences de l'ordre de 2×10^{-6} . Il a également été montré pour plusieurs de ces mutants que leur réversion est indépendante de la fonction de recA.

d) Mutations dans l'opérateur lac

Un autre groupe de mutations induites par la 2-aminopurine présentait également des résultats de recombinaison positifs avec des délétions telles que 5, 630 et 126, à des fréquences de 1 à 2×10^{-4} . Leur fréquence de réversion n'excédait cependant pas 10^{-8} . Ces résultats suggéraient donc qu'il s'agissait de mutations ponctuelles dans l'opérateur lac.

Il convient de mentionner ici que la souche utilisée pour la sélection de mutations I^- portait les mutations I^Q et L8 (promoteur lac) pour éviter que des mutants O^C partiellement constitutifs puissent croître sur le milieu de sélection contenant du Pgal (0,75 mg/ml) comme seule source de carbone. Afin d'expliquer l'apparition d'un groupe de mutations O^C aussi important, j'ai mesuré l'activité de β -galactosidase dans 3 de ces souches. Les valeurs observées dans les 3 cas étant identiques, 1 exemple est présenté au tableau 15, ainsi que les valeurs obtenues pour une mutation I^- dans une souche isogénique.

Il ressort de ces résultats que les mutations O^C en question ont provoqué une réversion partielle de L8.

Une influence mutuelle entre mutations de l'opérateur et du promoteur lac a été observée par Arditti et al. (1968) et Smith et Sadler (1971), mais jusqu'à présent aucune relation précise n'a pu être établie. La détermination du site (ou des sites) exact(s) de ces mutations permettrait éventuellement de mieux comprendre leur interaction.

D) DISCUSSION

1) Carte de délétions

L'analyse de plus de 400 délétions tonB-lacI isolées lors de ce travail a permis d'établir une carte du gène lacI de plus de 100 intervalles distincts. La localisation de ces délétions avec une grande collection de mutations I⁻ non-sense et missense a pu être effectuée rapidement grâce à une méthode de répllication et avec précision à l'aide d'une technique qui permet de détecter des recombinants à des fréquences de 10^{-7} à 10^{-8} . La résolution de cette carte représente une amélioration considérable par rapport aux cartes de délétions précédentes du gène I qui comprenaient au plus 20 intervalles (Ganem, 1972, Pfahl et al. 1974). Cependant l'avantage principal réside dans la corrélation entre cette carte génétique et la carte physique. En effet la majorité des mutations nonsense utilisées pour séparer les délétions ont été attribuées à des positions précises dans la protéine. (Miller et al. 1977, Coulondre et Miller, 1977a). Elles représentent par conséquent des limites précises pour les intervalles de délétions. Dans de nombreux cas cette corrélation a pu être étendue jusqu'aux nucléotides affectés par les mutations et l'élucidation récente de la séquence d'ADN du gène lacI a permis de la vérifier (P. Farabaugh, 1978). La connaissance des longueurs exactes de nombreux intervalles de délétions a permis d'effectuer plusieurs études qui seront commentées ci-dessous. De plus cette carte de délétions détaillée représente un des outils fondamentaux pour l'analyse de nombreux mutants de différents phénotypes (I^{-d}, I^s, I^{ts}) qui est actuellement effectuée pour tenter de localiser les sites actifs du répresseur lactose. Il est à noter également que la détermination directe de sites de mutations dans la séquence d'ADN du gène I est facilitée considérablement par leur localisation précise au moyen de cette carte.

2) Distribution des terminaisons de délétions dans le gène lacI

La distribution des terminaisons de délétions tonB-lacI à l'intérieur du gène I a été définie avec précision grâce aux nombreux marqueurs connus qui permettent d'examiner de petits intervalles (d'environ 15 nucléotides en moyenne). La comparaison avec une distribution de Poisson a révélé que la grande majorité des 400 délétions est répartie statistiquement le long du gène lacI, ce qui suggère que ces délétions définissent des segments plus nombreux et plus petits (d'une longueur moyenne de 3 à 5 nucléotides) que les intervalles démontrés. Cependant dans 2 ou 3 régions restreintes du gène I on constate un excès significatif de terminaisons de délétions par rapport à la distribution théorique. Il a été montré que des délétions se forment à certains sites préférentiels lorsqu'elles résultent d'une excision aberrante du phage λ (Pfeifer et al. 1974) ou d'éléments de translocation (voir revue par N. Kleckner, 1977) comme par exemple IS1 (Reif et Saedler, 1975, 1977). Les points chauds dans le gène lacI pourraient provenir d'un tel phénomène, mais leur cause réelle reste hypothétique, en particulier parce que l'autre extrémité des délétions située hors du gène I n'a pas été analysée avec précision.

3) Etudes de recombinaison

Les résultats obtenus lors de croisements entre différentes délétions et des mutations ponctuelles du gène lacI ont conduit à deux constatations principales:

- 1) Lorsqu'une délétion est croisée avec plusieurs mutations ponctuelles situées à des distances connues, on observe que dans l'ensemble les fréquences de recombinants tendent à augmenter en fonction de la distance, mais les valeurs divergent considérablement d'une fonction linéaire. Ce phénomène nommé "effet de marqueur", a été remarqué lors de croisements entre mutations ponctuelles dans divers autres systèmes (Yanofsky et al. 1964, Drapeau et al. 1968,

Stadler et Kariya, 1973, Norkin, 1970, Moore et Sherman, 1975, 1977).

- ii) Les fréquences de recombinants mesurées dans des croisements entre une mutation ponctuelle et plusieurs délétions qui se terminent dans différentes régions connues du gène I semblent dépendre linéairement de la distance. Les valeurs absolues des fréquences pour une distance donnée varient cependant selon la mutation ponctuelle employée.

Ces observations indiquent clairement que des "effets de marqueur" influencent la recombinaison mutation - délétion mais que ces effets sont dus aux mutations ponctuelles. Par conséquent il est dangereux de déterminer l'ordre de différentes mutations ponctuelles à l'intérieur d'un même intervalle de délétions uniquement sur la base de fréquences de recombinants. Par contre, si les fréquences obtenues lors de croisements de plusieurs délétions avec un même marqueur présentent des différences significatives, on peut utiliser ces résultats pour subdiviser ce groupe de délétions.

Une étude particulière a été consacrée aux "effets de marqueur" causés par différentes substitutions de bases au même site. Les dérivés amber et ochre des 8 codons spécifiant la tyrosine (UAU ou UAC) qui ne diffèrent que par une seule base du codon original ont été testés avec des délétions proches. Dans cinq cas l'amber présente une fréquence de recombinants 5 à 10 fois supérieure à celle de l'ochre correspondant, et il a été montré que ces mutations provenaient des codons UAU. Le fait que les séquences voisines de ces codons ne se ressemblent pas et l'absence d'effet observé avec les dérivés des codons UAC suggèrent que c'est le mauvais appariement d'une paire de bases spécifique qui affecte la recombinaison de façon aussi frappante. Cette observation fournit une information intéressante pour la compréhension de certains "effets de marqueur", d'autant plus qu'il s'est avéré difficile de dégager une règle générale d'analyses similaires effectuées dans d'autres systèmes par

croisements de mutations ponctuelles. (Drapeau et al. 1968, Stadler et Kariya, 1973, Moore et Sherman, 1977).

4) Distribution de mutations I^- non suppressibles

822 mutations I^- , induites spontanément et par les mutagènes 2AP, mutT, NQ et UV, ont été localisées sur la carte de délétion détaillée. Différents tests ont révélé qu'à part les mutations missense, cette collection comportait une fraction importante de délétions partielles ou totales du gène lacI, ainsi que de mutations à haute fréquence de réversion (10^{-6}) et de mutations O^C . Récemment une analyse détaillée de la série de mutations spontanées par P. Farabaugh (non publié) a conduit à plusieurs découvertes intéressantes: Les 58 % des mutations I^- qui se distinguent par une fréquence de réversion de l'ordre de 10^{-6} à 10^{-5} proviennent de l'insertion d'une séquence de 4 bases dans une région où cette même séquence de 4 bases est répétée 3 fois dans la souche originale. Un autre point chaud (13 %) de mutations, classées comme ponctuelles par les tests de recombinaison, représente en fait la délétion d'une de ces séquences de 4 bases. Une grande partie des mutations définies lors de localisation comme délétions internes (13 %) a été séquencée. Dans plusieurs cas les segments délétés sont flanqués de séquences identiques de 8 bases aux deux extrémités. Le fait que certaines de ces délétions ont été isolées deux fois alors qu'il existe de nombreuses autres séquences de cette longueur répétées dans le gène I (communication personnelle de D. Galas) suggère qu'un mécanisme spécifique intervient dans leur formation. Par ailleurs deux mutations ont été identifiées comme insertions de séquences IS1 au début du gène I (M. Calos et L. Johnsrud, non publié). Il apparaît donc que la collection spontanée comprend tout au plus 16 % de mutations missense, et il est probable que parmi celles-ci quelques mutations frameshift (petites insertions ou délétions) n'ont pas été détectées.

Cette distribution étonnante ne semble pas due à un biais de sélection car l'analyse des mutations induites par les

différents mutagènes a révélé l'existence de nombreux sites où une mutation ponctuelle peut entraîner l'inactivation du répresseur lactose. La distribution de ces mutations missense, spécialement de celles induites par 2AP, est dominée par des points chauds qui reflètent très probablement des particularités de la séquence d'ADN favorables aux mécanismes d'action de ces mutagènes.

ANNEXE

E) ETUDE DU REPRESSEUR LAC MODIFIE PAR LA DELETION B116

Parmi les 9 délétions tonB-lacI qui semblent se terminer avant le gène de structure I, la délétion B116 a suscité un intérêt particulier car des tests de liaison avec l'IPTG ont montré qu'elle résultait en une forte surproduction de répresseur lactose. Ce répresseur a été isolé et l'analyse de sa séquence peptidique a révélé que la délétion B116 fusionne les régions d'ADN correspondant aux séquences "leader" des messagers lacI (Steege, 1977) et trp (Squires et al. 1976), et qu'elle conduit à la traduction d'une séquence qui n'est pas exprimée normalement.

1) La séquence "leader" de l'ARN messenger trp

A son extrémité 5' l'ARN messenger de l'opéron trp comprend 166 nucléotides qui précèdent le premier gène de structure (trpE) et qui sont désignés par le nom de "leader". (Squires et al. 1976) Aux environs du nucléotide 145 de cette séquence se trouve un site appelé "atténuateur" qui permet de régler l'expression de l'opéron trp indépendamment du système répresseur-opérateur. En effet, s'il y a excès de tryptophane, environ 85 % des messagers sont arrêtés à ce site, alors que dans des conditions de carence en tryptophane, ou lorsque "l'atténuateur" est délété, la majorité des molécules d'ARN polymérase dépassent cette région pour transcrire les gènes de structure de l'opéron. (Bertrand et al. 1976)

On a détecté un site de liaison des ribosomes au début de la séquence "leader". (Platt et al. 1976) Ce site contient un codon AUG (nucléotides 27-29) qui est suivi, après 13 codons spécifiant des acides aminés, par un codon nonsense UGA. (Yanofsky, non publié) Bien qu'un peptide correspondant à cette séquence n'ait pas encore été isolé, la possibilité demeure qu'il joue un rôle dans la régulation par "l'atténuateur" de l'expression de l'opéron trp.

Ce peptide hypothétique pourrait induire ou inhiber le mécanisme de terminaison à "l'atténuateur" qui semble dépendre, (au moins partiellement) du facteur rho (Korn et Yanofsky, 1976).

2) Résultats

Les dosages présentés au tableau 1 montrent que la délétion B116 résulte en une surproduction de répresseur dans une souche trpR⁺ et que sous l'effet de la mutation trpR⁻ ce niveau de répresseur devient environ 150 fois plus élevé que dans une souche type sauvage. Compte tenu de cette surproduction extraordinaire, il semblait paradoxal que les niveaux résiduels de β -galactosidase fussent supérieurs à celui du type sauvage et surtout que ce niveau fût plus élevé dans la souche trpR⁻. Des tests en présence d'épisomes portant les allèles I⁺Z⁺ et I⁺Z⁻ ont montré qu'à l'état trpR⁺, l'activité du répresseur produit par B116 est insuffisante pour assurer une répression complète, alors que dans la souche trpR⁻, la synthèse accrue de β -galactosidase est insensible à une répression en trans, suggérant qu'elle provient d'une transcription efficace au-delà du gène I. Ce phénomène de "read-through" avait déjà été remarqué auparavant avec d'autres délétions trp-lac (Reznikoff et al. 1969, Mitchell et al. 1975, 1976).

Par ailleurs il a été montré que le produit de la délétion B116 immunoprécipité (par l'anticorps contre le répresseur sauvage) migre plus lentement sur gel de polyacrylamide en présence de SDS que le répresseur sauvage, ce qui indique une différence de taille estimée à environ 20 acides aminés. (A. Schmitz). L'observation que la protéine B116 est synthétisée à un taux si élevé et qu'elle est plus grande que le répresseur lac suggérerait fortement que la délétion B116 délète "l'atténuateur" de la séquence "leader" trp et que la traduction commence au site de liaison des ribosomes situé à l'extrémité 5' de ce messenger.

Cette hypothèse a été confirmée par l'analyse des acides

aminés de la partie amino-terminale du produit de B116 purifié (H. Sommer, H.G. Wittmann): les 7 premiers acides aminés correspondent à la séquence du leader trp qui suit immédiatement l'AUG du site de liaison des ribosomes. Les 9 résidus suivants qui précèdent ceux du répresseur sauvage sont codés par la séquence leader du gène lacI (Steege, 1977). Quant au codon GUG en position 17 qui spécifie normalement la méthionine initiale du répresseur lac, il est traduit comme valine à l'intérieur de la protéine B116. Au-delà de cette position la séquence du répresseur lac sauvage a été retrouvée. (voir figure 17)

Cette prolongation de la partie amino-terminale affecte les propriétés d'aggrégation du répresseur car divers polymères (en majorité des dodécamères) ont été détectés dans une préparation de répresseur B116 pur (soit par chromatographie sur colonne de Sephadex, H. Sommer, soit par analyse au microscope électronique, E. Boy de la Tour). Ces agrégats présentent une affinité accrue pour l'ADN non-opérateur alors que leur activité de liaison à l'opérateur est fortement réduite. L'activité résiduelle est éventuellement due à une faible proportion de tetramères. (A. Schmitz)

3) Conclusions

L'analyse détaillée du produit de fusion trp-lacI B116 a révélé que le site de liaison des ribosomes dans la séquence "leader" trp peut servir de site d'initiation pour la traduction in vivo. Bien que la délétion B116 crée une situation particulière, la forte synthèse de répresseur B116 suggère que ce site est également utilisé dans une souche de type sauvage. Le rôle éventuel soit du processus de traduction soit du produit fini dans la régulation de la transcription de l'opéron trp reste cependant encore à élucider. (Lee et Yanofsky, 1977).

L'extrémité opposée de la délétion B116 se termine exactement avant les 28 nucléotides de la région "leader" du

gène lacI, ce qui entraîne la traduction d'un fragment d'ARN qui n'est pas exprimé normalement. La séquence en acides aminés spécifiée par cette région a d'ailleurs permis de résoudre une ambiguïté dans la séquence de nucléotides déterminée par D. Steege (1977) (voir figure 17). Il est intéressant de noter aussi que dans les conditions créées par la délétion B116, le site d'initiation pour la traduction du répresseur lactose semble complètement masqué car on n'a pas détecté de répresseur type sauvage (moins de 1%). Cette inhibition peut être due à un changement de conformation de l'ARN messager, et/ou à une interférence avec l'appareil de traduction.

LEGENDE DES TABLEAUX

Tableau 1

Activités de β -galactosidase et de répresseur dans les souches contenant des délétions qui se terminent dans la région de contrôle du gène lacI. Les dosages ont été effectués selon les procédés décrits par Miller (1972). Les valeurs indiquées représentent la moyenne de plusieurs déterminations indépendantes.

La souche X7800 (Q) sert de référence pour les activités de répresseur. Elle porte la mutation I^Q qui permet une production dix fois plus élevée de répresseur que dans le type sauvage. (Müller-Hill et al. 1968)

Tableau 2

Les 408 souches de délétions tonB - lacI ont été testées par répllication sur milieu minimal glucose + proline avec et sans tryptophane. Les souches capables de croître sans tryptophane exogène sont désignées par le signe +.

Tableau 3

Fréquences de recombinants observées dans des croisements entre plusieurs mutations ponctuelles et des délétions qui se terminent près d'elles. Les valeurs indiquées représentent la moyenne de 3 à 7 déterminations indépendantes, effectuées selon le procédé de localisation à haute résolution (voir texte et matériel et méthodes).

Les positions des mutations ponctuelles ont été définies par analyse peptidique des répresseurs altérés (pour les résidus 5, 7, 16, 17, 19, Weber, 1975) ou par la méthode de corrélation mise au point par J. Miller, (pour les résidus 7, 17 et 18, Miller et al. 1977, pour le résidu 6, Coulondre et Miller, 1977a).

Tableau 4

Quelques basses fréquences de recombinants obtenues lors de croisements entre délétions et mutations ponctuelles ont été vérifiées par le test de réversion décrit dans Matériel et Méthodes. Pour chaque mutation ponctuelle la fréquence de révertants a été déterminée avec une délétion qui couvre la mutation en question. Les nombres de cellules totales et de colonies I^+ indiqués représentent les résultats obtenus avec au moins trois cultures indépendantes dans chaque cas.

Tableau 5

Corrélation des mutants nonsense utilisés pour la construction de la carte de délétions avec les résidus qu'ils affectent dans le répresseur (détails voir Miller et al. 1977, Coulondre & Miller, 1977a). Les sites amber, ochre ou UGA pour lesquels il est certain qu'une seule base diffère du codon original sont désignés par A, O ou U respectivement. Quant aux sites amber et ochre soit mal induits par les divers mutagènes, soit créés par substitutions de bases multiples, ils sont désignés par X et Y respectivement.

Pour la majorité des sites, les corrélations ont pu être vérifiées ou précisées grâce à la détermination récente de la séquence d'ADN du gène lac I (Farabaugh 1978). Certains de ces marqueurs ont été vérifiés indépendamment par analyse peptidique des répresseurs altérés. (cf. légende de fig. 8).

Tableau 6

Fréquences de recombinants typiques observées lors de croisements entre une mutation et différentes délétions ton B - lac I comprises dans un même intervalle démontré.

Tableau 7

Fréquences de recombinants voisines de la limite de résolution vérifiées par le test de réversion comme décrit dans Matériel et Méthodes (voir aussi tableau 4).

Tableau 8

Le segment de 1098 nucléotides compris entre la mutation O1 (Coulondre & Miller, 1977a) et L8 (Maxam & Sundaresan, non publié) a été divisé en 51 intervalles par des marqueurs connus. Ceux-ci ont été déterminés soit par la méthode de corrélation mise au point par J. Miller (tabl. 5) soit par analyse peptidique du répresseur ou de séquence d'ADN du gène lac I (pour références voir légende fig. 8). Le nombre de terminaisons de délétions trouvées dans chacun de ces groupes a été normalisé à des intervalles réguliers de 15 nucléotides. La moyenne de délétions par intervalle obtenue à partir des valeurs normalisées a été utilisée pour calculer la distribution de Poisson représentée à la figure 9.

Tableau 8bis

Pour les 4 intervalles dont le nombre de délétions semble, après normalisation, dépasser les valeurs théoriques (fig. 9), des distributions de Poisson ont été calculées sur la base des nombres observés. Les valeurs de $1 - \sum_{n=0}^{k-1} P$ représentent la probabilité de trouver un intervalle comprenant le nombre de terminaisons de délétions en question (k) ou supérieur.

Tableau 9

La délétion B 18 qui semble se terminer près du codon spécifiant le résidu 5 de la protéine, a été croisée avec 20 mutations nonsense connues, situées à des distances croissantes dans le gène lac I. Les fréquences de recombinants indiquées représentent la moyenne de 3 à 10 déterminations indépendantes. Dans deux cas les dérivés amber et ochre du même codon ont été testés.

Tableau 10

Les dérivés ochre et amber de certains codons spécifiant la leucine, la glutamine et l'acide glutamique ont été testés avec des délétions proches. Les fréquences de recombinants sont les moyennes d'au moins 3 déterminations indépendantes dans chaque cas.

Tableau 11

Croisements entre les dérivés ochre et amber des codons spécifiant la tyrosine et des délétions proches. Dans chaque cas 3 à 9 croisements indépendants ont été effectués pour déterminer la fréquence de recombinants.

Tableau 12

Résultats de croisements entre les dérivés ochre et amber du codon spécifiant la tyrosine en position 126, (Miller et al. 1977), et une série de délétions à distances croissantes. Les sites de terminaison approximatifs des délétions sont indiqués par les marqueurs connus qu'elles séparent. Les fréquences de recombinants représentent la moyenne de 3 à 6 déterminations indépendantes. L'analyse des recombinants après 15 générations supplémentaires a montré que pour les deux diploïdes avec la délétion 116 le rapport des fréquences de recombinants UAG/UAA reste constant.

Tableau 13

Répartition des 825 mutations I⁻ non suppressibles dans les différentes classes détectées par les analyses décrites dans le texte.

Récemment P. Farabaugh a montré que 18 des mutations spontanées définies comme ponctuelles par les techniques utilisées dans ce travail, sont en fait des délétions de 4 paires de bases.

Tableau 14

Liste des mutants dont la fréquence de réversion est égale ou supérieure à 10^{-6} .

Les 80 mutants spontanés de cette catégorie (voir tableau 13) n'ont pas été conservés.

Tableau 15

Activité de β -galactosidase mesurée dans 3 des 38 mutants O^C qui ont été isolés après traitement avec 2AP. Le contrôle est un mutant I⁻ de la même provenance.

LEGENDE DES FIGURES

Figure 1

Représentation schématique de la région ton B - trp du chromosome de la souche X7800. L'opéron lactose a été transposé au site att80 par un phage transducteur $\phi 80dlac$ dont il remplace des gènes tardifs. (Signer, Beckwith, 1966). Sous ce schéma sont indiqués les trois types de délétions résultant en une inactivation complète ou partielle du gène lac I, qui sont décrits dans le texte.

Figure 2

Mutations et délétions utilisées pour illustrer les limites de résolution de la méthode de localisation à haute résolution. (voir tableau 3). Les positions précises des mutations affectant les résidus 5, 7, 16, 17 et 19 ont été déterminées par analyse peptidique. (Weber et al. 1975).

La substitution de base qui affecte l'acide aminé en position 6 a été déduite grâce à la méthode de corrélation (Coulondre & Miller, 1977a) et vérifiée par la détermination des séquences d'ADN (Farabaugh, 1978) et d'ARN (Steege, 1977) de cette région.

Figure 3

Illustration des résultats obtenus à l'aide de la méthode de localisation par répllication.

Une série de 52 mutants I⁻ a été croisée avec 6 délétions différentes par le procédé décrit dans le texte et Matériel et Méthodes.

Les délétions sont respectivement: a) 208 b) B28
c) 235 d) B41 e) 35 f) 22

Figure 4

- a) Disposition des 52 mutants I^- sur les boîtes présentées à la figure 3.
- b) Certaines corrélations avec les acides aminés affectés dans le répresseur.

Les mutants I^- ont été placés dans l'ordre des résidus affectés dans la protéine. Les positions précises des mutations correspondant aux résidus 5, 7, 12, 16 et 201 ont été déterminée par analyse peptidique (Weber et al. 1975, Sommer 1977) et les autres par la méthode de corrélation génétique. (Miller et al. 77, Coulondre & Miller, 78).

La comparaison de ce schéma avec les résultats de la figure 3 montre que la méthode de réplification permet de séparer des mutations distantes de

- a) 4 nucléotides (entre position 6 et 7)
- b) 15 nucléotides
- c) 10 nucléotides
- d) 13 nucléotides
- e) 59 nucléotides
- f) 27 nucléotides.

Figure 5 (a - c)

Regroupement des 408 délétions ton B - lac I dans 41 des intervalles définis par 52 mutations ponctuelles I^- . Il s'agit de 48 mutations amber (désignées par A ou X) et ochre (désignées par O ou Y) isolées par J. Miller et C. Coulondre, ainsi que de 4 mutations missense, à savoir 40, AP309, 738 et AP46 (Weber et al. 1972). Chaque délétion a été croisée avec la série de 52 mutations par la méthode de réplification, et par le procédé à haute résolution contre les deux mutations qui définissent l'intervalle correspondant. (Les procédés de localisation sont décrits dans le texte et dans Matériel et Méthodes).

Les numéros indiqués au-dessus des noms de mutants désignent leur position sur la grille utilisée pour la réplification (figure 4).

Le mutant originalement en position 21 sur cette grille a été éliminé sur la carte, car il résultait d'une double mutation, dont une identique à celle du mutant en position 20. (Communication personnelle de C. Coulondre).

Figure 6 (a-d)

Carte de délétions partielle obtenue par localisation de 84 mutations ponctuelles I⁻, comprenant 64 intervalles. La localisation (comme décrit auparavant) de 32 mutations nonsense supplémentaires (isolées par J. Miller et C. Coulondre) a permis de créer 23 nouveaux intervalles de délétions. Outre les mutants amber (désignés par A et X) et ochre (O et Y), quelques mutants UGA (désignés par U) ont été utilisés.

Chaque nouvel intervalle a été vérifié par la méthode de localisation à haute résolution.

Figure 7(a-b)

Carte de délétions détaillée du gène lac I.

Les résultats de mapping combinés des 84 mutations I⁻ de la carte précédente (fig 6) et de plus de 600 mutations I⁻ non suppressibles (en majorité missense) sont présentés. Les intervalles sont numérotés de 1 à 100 avec des sous-groupes désignés par A, B, etc. Les mutations indiquées dans des cadres ont été croisées avec toutes les délétions des intervalles adjacents par la méthode à haute résolution. Les autres mutations ont été testées contre toutes les délétions des 2 groupes voisins par réplification et par le procédé à haute résolution contre les premières délétions de ces groupes. Les noms de tous les mutants nonsense ont été soulignés. Les mutations missense ont été désignées, selon leur provenance mutagénique, par les lettres A (2AP), T (mut T), Q (NQO), U (UV) et S (spontané).

Figure 8 (a - b)

Corrélation de la carte de délétions détaillée avec la carte physique du répresseur.

Une grande partie des intervalles de délétions représentés à la figure 7 ont pu être définis en longueurs physiques grâce à la corrélation des sites nonsense avec des résidus précis dans la protéine (Miller et al. 77, Coulondre & Miller, 1977a).

(Fig. 8 (a - b) suite)

voir aussi tableau 5).

Les positions précises de certains marqueurs nonsense ont été déterminées indépendamment par analyse peptidique des répresseurs altérés, notamment aux résidus 7, 12, 17, 26, 60 (Weber et al, 1975) 61 et 62 (Files et al. 75) 201 et 220 (Sommer, 1977) et 3 (Schmitz et al. 1978). Quelques autres mutations ont été analysées directement sur la séquence d'ADN, à savoir celles affectant les résidus 269 (Farabaugh, non publié) 319 (Lillis, 1977) 311 et 351 (Maxam & Gilbert, non publié).

Pour les mutations missense, les seules positions déterminées par analyse peptidique sont 5, 16, 19 et 53 (Weber et al. 1975). Quant aux intervalles délimités par d'autres sites missense, leurs longueurs ont été définies proportionnellement au nombre de délétions qui s'y terminent.

N.B. Un sous-groupe supplémentaire a été créé dans l'intervalle 99 par localisation d'une mutation missense isolée par M. Hofer).

A cause d'un ajustement de la numérotation, le marqueur X10 de la figure 7 est indiqué ici sous le nom de X9.

Figure 9

Distribution des terminaisons de délétions dans le gène lac I.

L'histogramme représente la distribution expérimentale normalisée comme décrit dans le texte. (Les nombres normalisés de délétions par intervalle, présentés au tableau 8, ont été arrondis). La hauteur attribuée à chaque intervalle est de 0.02, de façon à ce que la somme des intervalles soit égale à 1.

La distribution de Poisson calculée pour la moyenne expérimentale de 6 terminaisons de délétions par intervalle est représentée par la courbe continue. (Bien que la distribution de Poisson soit une distribution discrète, cette représentation

(Fig. 9 suite)

a été choisie pour faciliter la comparaison avec les données expérimentales).

La ligne pointillée illustre la distribution de Poisson calculée après élimination des deux groupes dont le nombre (normalisé) de délétions observées dépasse largement les valeurs théoriques (intervalles à 17 et 20 terminaisons de délétions).

Figure 10

Représentation graphique des fréquences de recombinants en fonction de la distance entre une délétion (B18) et différentes mutations du gène lac I. (Les fréquences déterminées par 3 à 10 croisements indépendants, ainsi que les écarts standard sont indiqués au tableau 9).

En abscisse figurent les positions de la protéine correspondant aux mutations utilisées. Dans deux cas, les dérivés amber et ochre du même codon ont été testés, ce qui est indiqué par A et O respectivement.

Figure 11

Fréquences de recombinants en fonction de la distance qui sépare différentes délétions d'une mutation ponctuelle.

Les mutations amber (cercles) et ochre (points) dérivées du codon spécifiant la tyrosine 126 (Miller et al. 1977) ont été croisées avec des délétions qui se terminent à différents endroits dans le gène lac I. En abscisse sont indiqués les marqueurs connus qui définissent les intervalles de délétions correspondants. Les droites ont été calculées selon la méthode des moindres carrés.

Figure 12

Localisation de 248 mutations missense I⁻ induites par la 2-amino-purine, (A), sur la carte de délétions détaillée.

(Fig. 12 suite)

Les mutations indiquées dans les cadres ont été croisées avec toutes les délétions des deux groupes limitrophes par la méthode à haute résolution. Les mutations dont les noms sont précédés d'un tiret ont été testées par répllication contre toutes les délétions et par la méthode à haute résolution contre les premières délétions des groupes voisins. Les autres mutations ont été localisées uniquement par répllication. A l'intérieur d'un intervalle de délétion, différents groupes de mutations ont été créés sur la base des résultats de complémentation négative (pour les mutations situées dans la région correspondant aux 60 - 70 premiers résidus) et de recombinaison avec des mutations ponctuelles dans l'autre partie du gène I (voir délétions internes). Dans deux cas, une mutation d'un groupe a été transférée sur le chromosome, puis testée contre tous les autres membres. (détails voir texte). Il s'agit de la série de mutations All - A30 et du groupe situé dans la région correspondant au résidu 251.

Figure 13 (a-h)

Localisation de 154 mutations missense I⁻ induites par la lumière ultraviolette (U) et de 78 induites par le 4-nitroquinoline-1-oxide. (Q)

Mêmes remarques générales que pour la figure 12. L'identité des 10 mutations (U) du dernier groupe a été vérifiée par recombinaison avec l'une d'elles (U197).

Figure 14 (a-h)

Localisation de 86 mutations missense I⁻ induites par le gène mutateur mutT (T) et de 41 spontanées (S). Mêmes remarques générales que pour la figure 12. Le groupe de 18 mutations spontanées (dans la région comprise entre les résidus 193 et 201) a été défini comme site unique par test de recombinaison. Pour 2 de ces mutations P. Farabaugh (non publié) a montré par analyse de séquence d'ADN qu'il s'agit d'une délétion de 4 bases.

Figure 14 (suite)

Il a été montré par ailleurs que les mutations S58 et S114 sont dues à l'insertion d'un élément IS1 au début du gène I. (M. Calos et L. Johnsrud, non publié)

Figure 15

Distribution des mutations missense représentée à l'échelle de la carte physique du répresseur (les résidus sont indiqués en abscisse) les positions respectives des mutations ont été déterminées approximativement à l'aide des marqueurs connus (voir cartes précédentes) et du nombre de délétions qui les séparent. Les mutations spontanées ne sont pas présentées, car le nombre de mutations ponctuelles ne représente qu'une partie minime de la collection (voir tableau 13).

Figure 16

Délétions partielles du gène lacI détectées parmi les mutations I⁻ non supprimibles (sur F'lacpro) par test de recombinaison avec une série de marqueurs connus sur le chromosome. (Les résidus affectés par ces derniers sont indiqués en abscisse.) Les sites de terminaison des délétions ont été déterminés par la méthode de localisation à haute résolution, décrite au chapitre Matériel et Méthodes.

Figure 17

Partie 5' de l'ARN messager et région amino-terminale de la protéine créés par la délétion B116. Les séquences "leader" trp (Bertrand et al. 1976) et lacI (Steege, 1977) fusionnées par la délétion B116 sont montrées, ainsi que les acides aminés correspondants de la protéine B116. L'incertitude sur la position de 6 bases de l'ARN messager (indiquées entre parenthèses) juste avant le codon d'initiation lacI, a été levée grâce à la séquence correspondante en acides aminés du répresseur B116.

TABLEAU 1

Élétion	trp R	Activité β -galactosidase sans IPTG	10^{-3} M IPTG	Rapport avec/sans IPTG	Activité spécif. liaison IPTG
196	+	667	642	1.0	-
196	-	517	547	1.1	-
101	+	102	586	5.8	< 0.05
101	-	0.8	234	293	0.25
119	+	244	724	3.0	< 0.05
119	-	15	379	25.3	0.05
B116	+	2.4	470	196	0.41
B116	-	11	168	15.3	13.5
43	+	155	588	3.8	< 0.05
43	-	15	214	14.3	0.38
27	+	210	539	2.6	< 0.05
27	-	54	513	9.5	< 0.05
B80	+	469	698	1.5	< 0.05
B80	-	327	608	1.9	0.05
B84	+	614	724	1.2	< 0.05
B84	-	481	553	1.1	< 0.05
B137	+	607	655	1.1	< 0.05
B137	-	580	723	1.2	< 0.05
245	+	595	608	1.0	< 0.05
245	-	561	598	1.1	< 0.05
X7800	+	0.5	416	832	-
X7800(Q)	+	1.0	257	257	0.90

TABLEAU 2

Caractère Trp des délétions TonB-LacI

Délétion	Trp	Délétion	Trp	Délétion	Trp	Délétion	Trp
1	-	35	+	70	+	105	+
2	+	36	-	71	-	106	-
3	-	38	-	72	-	107	-
4	-	39	-	73	-	108	+
5	-	40	-	74	-	109	-
6	-	41	-	75	-	110	-
7	+	42	+	76	+	111	-
8	+	43	-	77	-	112	-
9	-	44	+	78	-	113	-
10	-	45	-	79	-	114	+
11	-	46	-	80	-	115	-
12	+	47	+	81	-	116	-
13	-	48	-	83	-	117	+
14	+	49	+	84	-	118	-
15	-	50	+	85	-	119	-
16	+	51	-	86	+	120	+
17	-	52	-	87	+	121	-
18	-	53	+	88	-	122	-
19	-	54	-	89	-	123	-
20	+	55	+	90	+	124	+
21	+	56	-	91	-	125	-
22	-	57	+	92	+	126	+
23	+	58	-	93	-	127	-
24	-	59	+	94	+	128	-
25	+	60	+	95	+	129	-
26	-	61	-	96	-	130	-
27	+	62	-	97	+	131	-
28	+	63	+	98	-	132	-
29	-	64	+	99	+	133	+
30	-	65	-	100	-	134	-
31	+	66	-	101	-	135	+
32	+	67	-	102	-	136	-
33	+	68	-	103	-	137	-
34	-	69	-	104	+	138	+

TABLEAU 2 (suite)

Délétion	Trp	Délétion	Trp	Délétion	Trp	Délétion	Trp
139	-	174	-	209	+	244	+
140	+	175	-	210	-	245	+
141	+	176	-	211	-	246	-
142	-	177	-	212	+	247	+
143	-	178	-	213	-	248	+
144	+	179	+	214	+	249	-
145	+	180	+	215	+	250	-
146	+	181	-	216	+	B1	-
147	+	182	+	217	-	B3	+
148	-	183	+	218	+	B4	-
149	+	184	-	219	+	B5	+
150	-	185	-	220	-	B6	-
151	-	186	+	221	-	B8	-
152	+	187	-	222	-	B9	-
153	+	188	-	223	+	B10	-
154	-	189	-	224	+	B11	-
155	+	190	-	225	-	B12	-
156	-	191	-	226	-	B13	+
157	-	192	+	227	-	B14	+
158	-	193	+	228	+	B15	-
159	+	194	-	229	-	B16	-
160	+	195	-	230	+	B17	-
161	-	196	-	231	+	B18	-
162	-	197	+	232	+	B19	+
163	+	198	-	233	+	B20	-
164	+	199	-	234	+	B21	-
165	-	200	+	235	+	B22	+
166	+	201	-	236	-	B23	-
167	+	202	-	237	-	B24	-
168	-	203	-	238	-	B25	+
169	-	204	-	239	-	B26	+
170	-	205	-	240	+	B27	-
171	+	206	-	241	+	B28	-
172	-	207	-	242	+	B29	+
173	-	208	-	243	-	B30	-

TABLEAU 2 (fin)

Délétion	Trp	Délétion	Trp	Délétion	Trp	Délétion	Trp
B31	+	B70	-	B107	-	B150	+
B32	+	B71	+	B110	-	B151	+
B33	-	B72	+	B111	-	B152	-
B34	+	B73	-	B113	-	B153	-
B35	-	B74	-	B114	+	B154	-
B36	-	B75	+	B115	-	B155	-
B37	-	B76	+	B117	-	B156	-
B38	-	B77	+	B118	+	B157	+
B39	+	B78	+	B119	+	B158	+
B40	+	B79	+	B121	-	B159	+
B41	-	B80	+	B122	+	B160	+
B42	+	B81	+	B123	-	B161	+
B43	-	B82	-	B124	-	B162	-
B45	-	B83	-	B125	-	B163	+
B46	-	B84	+	B126	-	B164	-
B47	+	B85	-	B127	+	B165	-
B48	-	B86	+	B128	-	B166	+
B49	-	B87	-	B129	-	B167	+
B50	+	B88	-	B130	-	B168	+
B51	+	B89	-	B133	+	B169	-
B53	-	B90	-	B134	-	B170	-
B54	-	B91	+	B135	+	B171	+
B55	+	B92	+	B136	-	B172	-
B56	-	B94	-	B137	-	B173	+
B57	+	B95	-	B138	+	B174	-
B58	-	B96	+	B139	+	B175	+
B59	-	B97	+	B140	-	B176	-
B60	+	B98	+	B141	+		
B61	-	B99	-	B142	-	B116	-
B63	-	B100	+	B143	-		
B64	-	B101	+	B144	-		
B65	-	B103	+	B146	-		
B66	-	B104	-	B147	-		
B67	+	B105	+	B148	-		
B69	+	B106	+	B149	-		

TABLEAU 3

Tests de recombinaison avec des marqueurs proches

Résidu affecté par mutation ponctuelle	Nucléotide affecté	Délétion	Fréquence de recombinants (10^{-7})
5	14	208	< 0.1
6	17	208	2.5
6	17	24	< 0.1
7	21	24	4.3
7	21	B28	< 0.1
16	46	B13	< 0.1
17	51	B13	14.4
18	52	B13	13.5
19	55	B13	8.6
17	51	129	< 0.1
18	52	129	< 0.1
19	55	129	< 0.1

TABLEAU 4

Tests de réversion

Mutation	Délétion	Cell.totales	Colonies I ⁺	Fréquence de recombinants (10 ⁻⁸)
U180	B33	6.9 x 10 ⁸	110	16
U180	85	7.5 x 10 ⁸	0	< 1
O30	86	6.0 x 10 ⁸	59	9.8
O30	22	1.0 x 10 ⁹	0	< 1
A36	33	1.2 x 10 ⁹	155	13
A36	B50	1.0 x 10 ⁹	6	< 1
U56	155	1.3 x 10 ⁹	124	9.8
U56	B40	1.3 x 10 ⁹	1	< 1

T A B L E A U 5

Corrélation des sites nonsense

Mutant	Résidu affecté	Mutant	Résidu affecté	Mutant	Résidu affecté
A1	Tyr 7	01	Lys 2	X1	Thr 5
A2	Glu 11	02	Leu 6	X2	Leu 45
A3	Tyr 12	06	?	X3	Leu 90?
A4	Tyr 17	07	?	X4	Leu 122
A5	Gln 18	08	Tyr 47	X5	Leu 128
A6	Gln 26	09	Gln 55	X6	Leu 184
A7	Glu 44	010	Gln 54	X7	Leu 251
A9	Gln 60	011	Gln 78	X8	Leu 295/6
A10	Ser 61	012	Lys 84	X10	Leu 323
A11	Leu 62	013	Gln 89	X13	Leu 318
A12	Ser 77	014	Glu 100	Y1	Pro 3
A13	Ser 97	015	Glu 105	Y2	Lys 59
A14	Tyr 126	016	Lys 108	Y3	Leu 146
A15	Gln 131	017	Gln 117	Y4	Lys 203?
A16	Gln 153	018	Tyr 126	Y6	Lys 290?
A17	Glu 172?	019	Glu 137	Y8	Lys 325/7?
A18	Leu 177	020	Glu 164	U1	Arg 101
A19	Gln 180	021	Gln 181	U2	Cys 140
A20	Ser 193	022	Leu 189	U3	Cys 107
A21	Trp 201	024	Gln 209	U8	Cys 281
A22	Tyr 204	025	Glu 215	U10	Arg 168
023	Gln 211	026	Glu 217		
A24	Trp 220	027	Gln 227		
A25	Glu 235	028	Gln 228		
A26	Gln 248	029	Gln 231		
A27	Glu 259	030	Tyr 273		
A28	Ser 269	031	Glu 277		
A30	Tyr 282	032	Ser 280		
A31	Gln 291	034	Gln 298		
A32	Leu 304	035	Gln 206		
A33	Gln 309	036	Ser 322		
A34	Gln 311				
A35	Gln 317				
A36	Leu 319				

TABLEAU 6

Groupe de Délétions	Délétion Chromosome	Mutation Episome	Fréquence de Recombinants (10^{-7})
99A	137	L1	100
99B	21	L1	14
99C	59	L1	0.8
100A	5	L1	0.1
2A	196	40	60
2B	B18	40	2.8
3	208	40	0.1

TABLEAU 7

Tests de Réversion

Groupe de Délétions	Délétion	Mutation	Cellules totales	Colonies I ⁺	Fréq. Recomb. (10 ⁻⁸)
29A	46	X3	1.2 10 ⁹	33	2.7
29B	61	X3	1.0 10 ⁹	2	<1
96A	B147	Q20	1.2 10 ⁹	57	4.8
96B	33	Q20	1.4 10 ⁹	0	<1
38B	190	A14	3.8 10 ⁸	14	3.7
38B	190	O18	4.0 10 ⁸	0	<1
39	B41	A14	4.2 10 ⁸	0	<1

TABLEAU 8

Distribution des terminaisons de délétions dans le gène lacI

Intervalle (nucléotides spécifiés par marqueurs)	Longueur de l'intervalle (nucléotides)	Nombre de délétions par intervalle	Nombre de délétions normalisé à intervalle de 15 nucléotides
4-21	17	9	7.9
21-36	15	3	3.0
36-51	15	6	6.0
51-76	25	6	3.6
76-130	54	17	4.7
130-141	11	4	5.5
141-163	22	7	4.8
163-185	22	5	3.4
185-230	45	10	3.3
230-250	20	5	3.8
250-268/9	18	4	3.3
268/9-290	21	7	5.0
290-301	11	5	6.8
301-313	12	1	1.2
313-322	9	6	10.0
322-349	27	7	3.9
349-364/5	15	4	4.0
364/5-378	13	4	4.6
378-391	13	2	2.3
391-420	29	5	2.6
420-437	17	8	7.1
437-457	20	9	6.8
457-490	33	10	4.6
490-514	24	12	7.5
514-530	16	8	7.5
530-541	11	2	2.7
541-550/1	9	3	5.0
550/1-566	15	6	6.0
566-578	12	6	7.5
578-602	24	5	3.1
602-625	23	1	0.7

TABLEAU 8 (suite)

Intervalle (nucléotides spécifiés par marqueurs)	Longueur de l'intervalle (nucléotides)	Nombre de délétions par intervalle	Nombre de délétions normalisé à intervalle de 15 nucléotides
625-643	18	13	10.8
643-659	16	1	0.9
659-679	20	8	6.0
679-691	12	6	7.5
691-703	12	9	11.3
703-742	39	9	3.5
742-751/2	9	8	13.3
751/2-775	23	7	4.6
775-806	31	12	5.8
806-819	13	17	19.6
819-829	10	11	16.5
829-846	17	6	5.3
846-871	25	6	3.6
871-892	21	19	13.6
892-916	24	8	5.0
916-931	15	6	6.0
931-949	18	12	10.0
949-965	16	7	6.6
965-1052	92	26	4.5
1052-1102	45	18	5.4
4 - 1102	1098	396	6.0

TABLEAU 8 bis

Groupe (cf.fig.9)	Longueur de l'intervalle (nucléotides)	Nombre moyen de délétions par intervalle	Nombre observé de délétions dans intervalle (k)	$1 - \sum_{n=0}^{k-1} P$
I	9	3.2	8	0.016
II	21	7.6	19	0.0004
III	10	3.6	11	0.001
IV	13	4.7	17	$9 \cdot 10^{-6}$

TABLEAU 9

Fréquences de recombinants observées lors de croisements entre
la délétion B18 et différents marqueurs

Résidu affecté		Fréquence de recombinants (10^{-7})	
5		2.8 ±	0.8
6		7.8 ±	1.3
7		22 ±	4
16		52 ±	5
26		545 ±	50
47		755 ±	90
60		655 ±	120
78		1120 ±	40
89		1250 ±	280
100		1950 ±	160
101		820 ±	60
117		1130 ±	220
126	amber	2200 ±	320
126	ochre	1150 ±	40
153		1780 ±	180
201		1880 ±	230
220		3400 ±	250
248		2050 ±	70
282	amber	3220 ±	220
282	ochre	3030 ±	310

TABLEAU 10

Mutations ochre et amber dérivées des codons UUA, CAA et GAA

Résidu	Codon type sauvage	Délétion	Fréquence de recombinaants (10^{-7})	
			UAA	UAG
Leu 6	UUA	B18	5.1	2.5
Gln 78	CAA	64	12.8	2.2
Gln 89	CAA	191	22.2	15.9
Glu 105	GAA	B138	24.4	8.9
Glu 137	GAA	12	9.1	7.1
Glu 215	GAA	189	15.0	3.3
Gln 227	CAA	142	8.3	15.8

TABLEAU 11

Mutations ochre et amber dérivées des codons de la tyrosine

Positions des tyrosines	Codon type sauvage	Délétion	Fréquence de recombinaants (10^{-7})	
			UAG	UAA
7	UAC	208	6.5 $\pm$ 3.6	5.8 $\pm$ 1.1
47	UAC	56	4.3 $\pm$ 0.8	4.6 $\pm$ 0.6
273	UAC	4	25 $\pm$ 2	23 $\pm$ 2
12	UAU	26	45 $\pm$ 10	5.1 $\pm$ 0.1
17	UAU	235	42 $\pm$ 12	6.5 $\pm$ 1.0
126	UAU	116	72 $\pm$ 6	7.1 $\pm$ 1.2
204	UAU	B76	8.5 $\pm$ 2.4	1.1 $\pm$ 0.3
282	UAU	3	27 $\pm$ 1	6.3 $\pm$ 0.6

TABLEAU 12

Recombinaison de deux mutations au même site
avec différentes délétions

Délétion	Terminaison approximative (résidu)	Fréquence (10^{-7})		Rapport UAG/UAA
		UAG	UAA	
116	123-126	70 $\pm$ 8	7.2 $\pm$ 1	9.7
94	118-122	280 $\pm$ 16	67 $\pm$ 4	4.2
77	117	240 $\pm$ 28	71 $\pm$ 8	3.4
20	110	400 $\pm$ 44	130 $\pm$ 6	3.1
156	95-100	420 $\pm$ 35	280 $\pm$ 57	1.5
46	90-95	670 $\pm$ 58	320 $\pm$ 8	2.1
25	63-75	900 $\pm$ 50	790 $\pm$ 69	1.1
210	48-53	1300 $\pm$ 150	600 $\pm$ 48	2.2
196	3-5	1800 $\pm$ 190	1000 $\pm$ 66	1.8

TABLEAU 13

Mutations I⁻ non supprimables

Mutagène	a) Mutations ponctuelles	b) Délétions internes	c) Mutations à haute fréq. réversion	d) Mutations O ^c	Nombre total de mutations
2AP	248	9	40	38	335
UV	154	11	11		176
NQO	78	1	8		87
mutT	86		2		88
spontané	40	19	80		139
Total	606	40	141	38	825

TABLEAU 14

Mutations à haute fréquence de réversion

Mutagène	Noms des mutants
2AP	A117 - A124, A126 - A135 A137 - A158
NQO	Q12, Q16, Q18, Q26, Q70, Q75, Q78, Q84
UV	U51, U52, U59, U71, U73, U75 U77, U78, U131, U139, U181
mut T	T61, T105

TABLEAU 15

Analyse de mutants O^c

Mutant	Génotype	Activité β -galactosidase	
		-IPTG	+IPTG
ANC101	O ^c	179	471
A 153	I ⁻	106	102

X7800

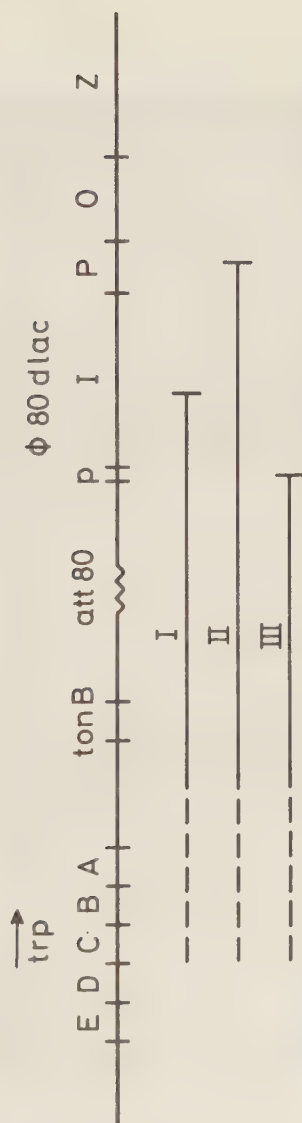


Figure 1

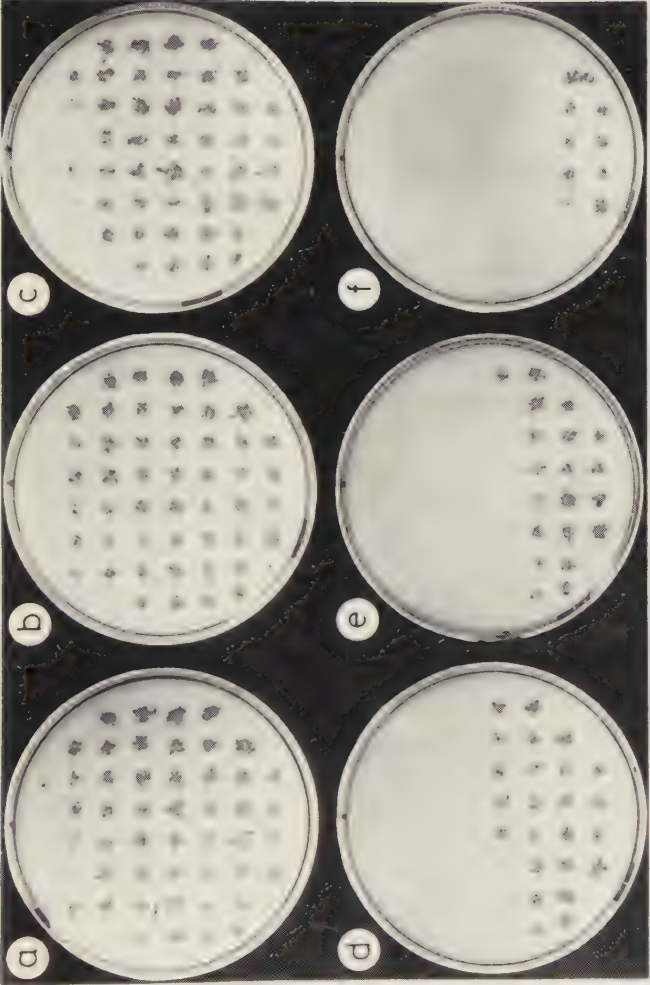
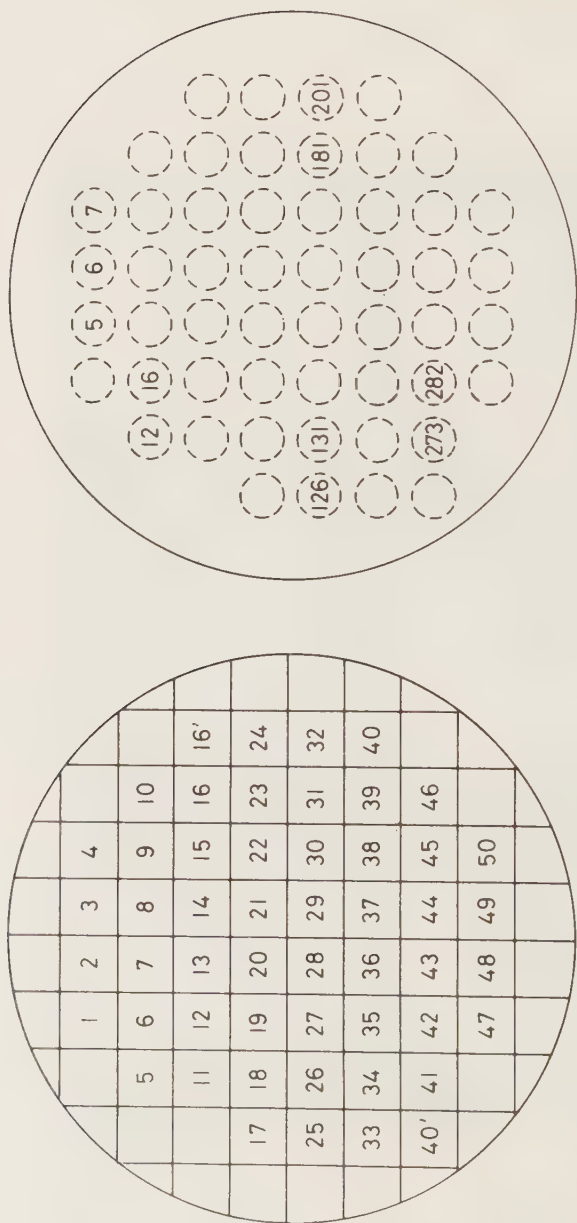


Figure 3



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	16.5
01	40	02	A1	A3	AP309	A4	A5	738	A6	A7	AP46	010	Y2	A9	A10	All
27 43 101 119 245 B80 B84 B137 B116	196 B18 208 B29 B111	1 24 70 B99	26 B28 B134	235 B49	247 B13 B22 B94			129 130 140 B124	2 8 16 23 44 49 81 92 111 112 143 147 188 242 B56	19 56 102 105 163 178 210 B43 B45 B57 B63 B95 B105 B141 B150		109 118 128 B162 B170				25 53 64 113 184 B5 B47 B78 B135 B168

Figure 5a

17	18	19	20	22	23	24	25	26	27	28	29	30	31	32	33	34
A12	O11	O13	A13	O14	O17	X4	O18	A15	O19	A16	A18	A19	O21	A21	A22	A23
76 191 197 230 244 B110 B143		46 61 100 131 B32 B117 B138 B149		20 47 57 77 97 103 115 123 127 156 217 229 237 241 243 B26 B42 B71 B81 B115	94 108 223 B6	116 200 239 B3	190 B41		6 7 12 17 30 31 114 133 149 150 162 209 B4 B11 B23 B54 B65 B79 B104 B107 B122 B125 B129 B166	10 68 69 17 90 93 121 135 141 161 174 179 182 202 215 226 234 B10 B51 B61 B70 B75 B103 B146 B157	106 221 B153 B154 B163		32 35 55 73 151 155 172 211 220 222 246 B21 B40 B53 B59 B77 B90 B118 B136 B158	B76	40 84 B33	85 124 167 189 193 240 B39 B66 B92 B133 B160

Figure 5b

35	36	37	38	39	40	40.5	41	42	43	44	45	46	47	48	49	50
A24	Q27	Q28	Q29	A26	A27	A28	Q30	A30	A31	Q34	Q35	A33	A34	A35	A36	Q36
142		75	34	50	9	4	3	54	13	42	51	218	11	33	B50	5
169		78	144	63	58	45	22	120	39	91	88		29	136	B60	14
177		83	148	67	98	52	38	157	41	117	204		89	B147	B121	15
185		87	165	95	104	60	62	164	66	122	219		159	B148		18
207		107	168	132	194	170	71	166	125	233	B67		160			21
216		153	187	139	B14	203	72	180	138	B46			175			28
B19			236	250	B15	224	86	B119	158	B98			186			36
B126			238	B1	B20	B9	96		181	B167			195			48
			B36	B30	B25	B34	134		192				231			59
			B69	B64	B40	B161	173		227				B17			65
			B72	B73	B82	B171	198		B8				B155			74
			B74	B106	B91		199		B12							79
			B85	B128	B123		205		B31							99
			B89	B151			212		B48							110
			B113	B164			213		B83							126
			B114				249		B127							137
			B140				B16		B144							B100
			B142				B35		B165							145
							B96									146
							B139									B130
							B156									152
							B172									B152
							B173									154
																B159
																176
																B169
																183
																B174
																201
																B175
																206
																B176

Figure 5c

	Y1	X1	O2	A1	A3	A2	A309	A4	A5	738	O6	O7	A7	X2	O8	C9	O10	A9	O11	O12
	O1	40									A6					AP46		Y2	A12	O12
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17				
27	196	208	1	26	235	B13	111	8	2	19	56	102	109	25	76	191				
43	B18	B29	24	B28	B49	B22	129	16	49		105	163	118	53	197	B110				
101		B111	70	B134		B94	130	23	81		B105	178	128	64	230					
119			B99			247	140	44	92			210	B162	113	244					
245							B56	147	112			B57	B170	184	B143					
B80							B141	188	143			B63		B5						
B84							B150	242	242			B95		B47						
B137								B43	B43					B78						
B116								B45	B45					B137						
								B141	B141					B168						

Figure 6a

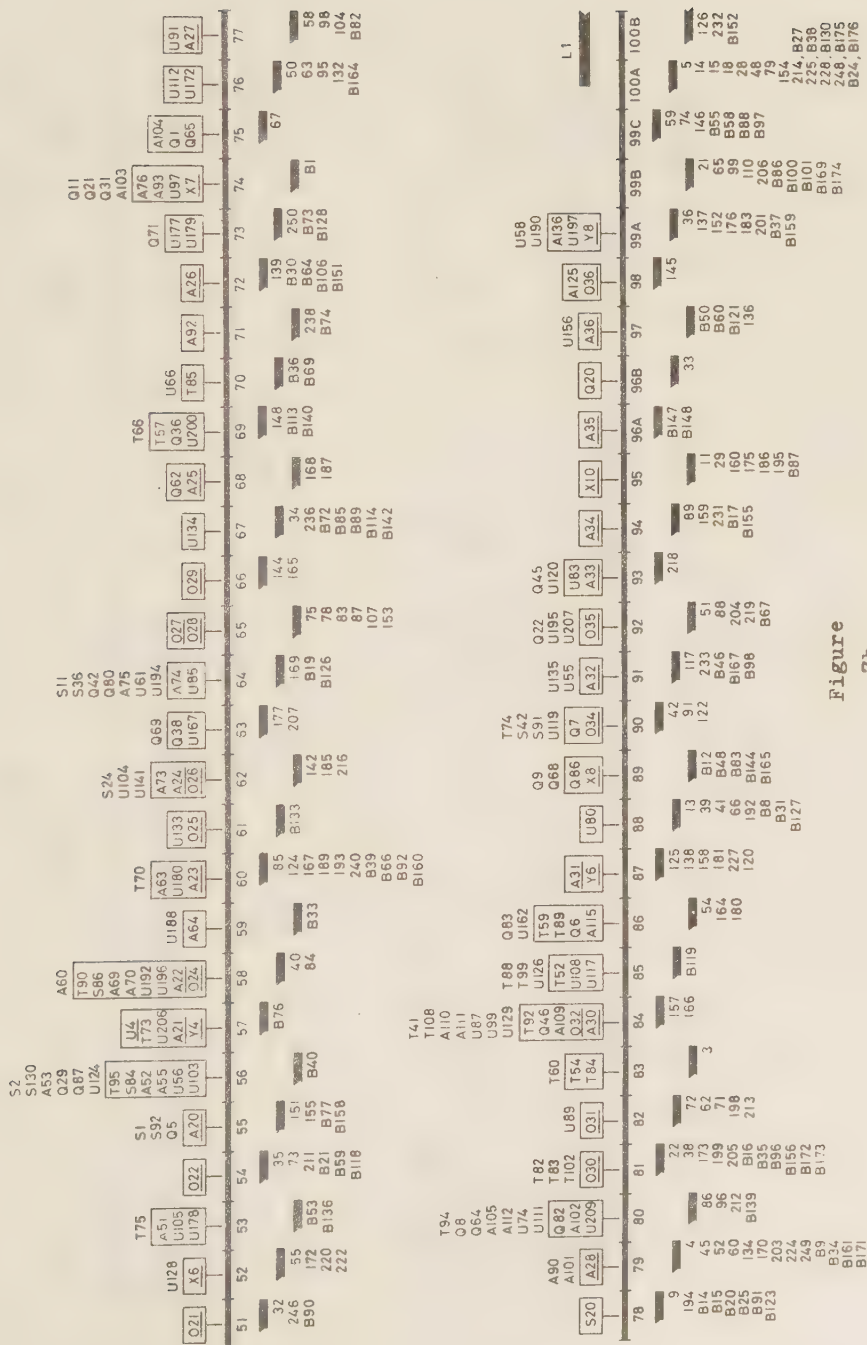
X3	U1	U16	U2	Y3	A16	O20	A17	A18	O21
O13	A13	O14	O15	U3	X5	A15	O19		A19
18	19	20	21	22	23	24	25	26	27
>	>	>	>	>	>	>	>	>	>
46	156	103	47	20	94	116	190	B41	12
61	241		115	57	108	200			17
100	243		127	77	223	239			162
123	B26		237	97	B6	B3			209
131	B81		B42	217					B122
B32			B71	229					B104
B138			B115						B125
B149									B129
									B107
									B166
									B79
									B61
									B23
									B25
									B103
									B157
									B65
									B70
									B146
									B154
									B153
									B21
									B163
									246
									B90
									32
									10
									93
									106
									161
									202
									179
									141
									90
									80
									68
									69
									121
									135
									182
									174
									121
									141
									174
									179
									202
									161
									106
									93
									10
									32
									33
									34

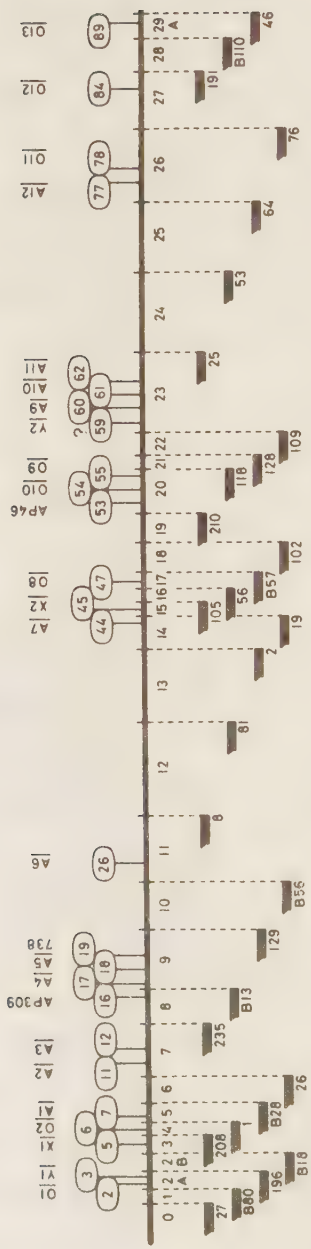
Figure 6b

X6	O22	A20	Y4	A21	O24	A22	A23	O25	A24	O26	O27	O28	A25	A26	X7	A27	A28	O30
35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50			
>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>
55	35	151	B76	40	85	B133	142	75	34	148	139	50	9	4	22			
172	73	155	84	B33	124	124	169	78	144	168	250	63	58	45	38			
220	211	B40			167	177	177	83	165	187	B30	67	98	52	173			
222	B21	B77			189	185	185	87	236	238	B64	95	104	60	199			
B53	B59	B158			193	207	207	107	B72	B36	B73	132	194	86	205			
B136	B118				240	216	216	153	B85	B69	B106	B1	B14	96	B16			
					B39	B19	B19		B89	B74	B128	B164	B15	134	B35			
					B66	B126	B126		B114	B113	B151		B20	170	B96			
					B92				B142	B140			B25	203	B156			
					B160								B82	212	B172			
													B91	224	B173			
													B123	249				
														B9				
														B34				
														B139				
														B161				
														B171				

Figure 6c







Figure

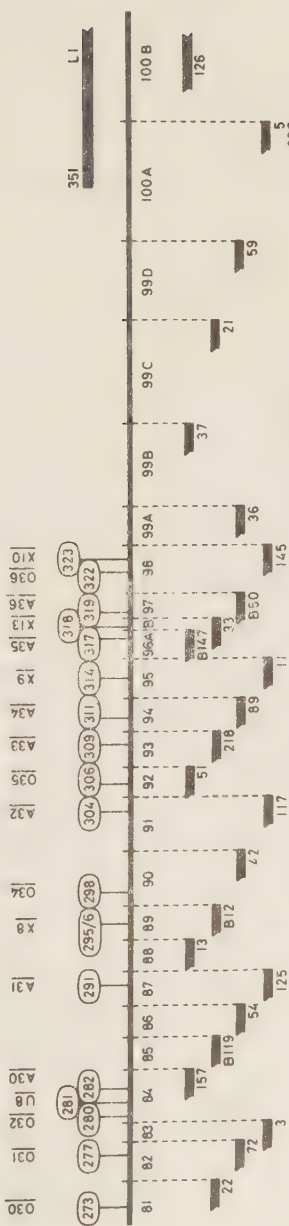
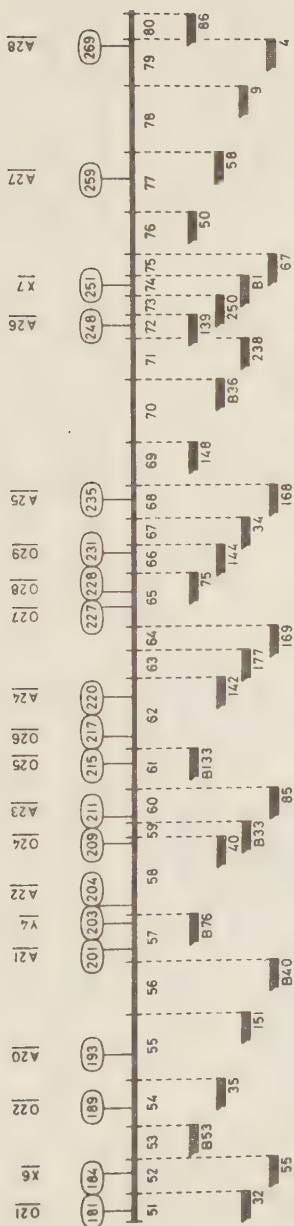


Figure 8b

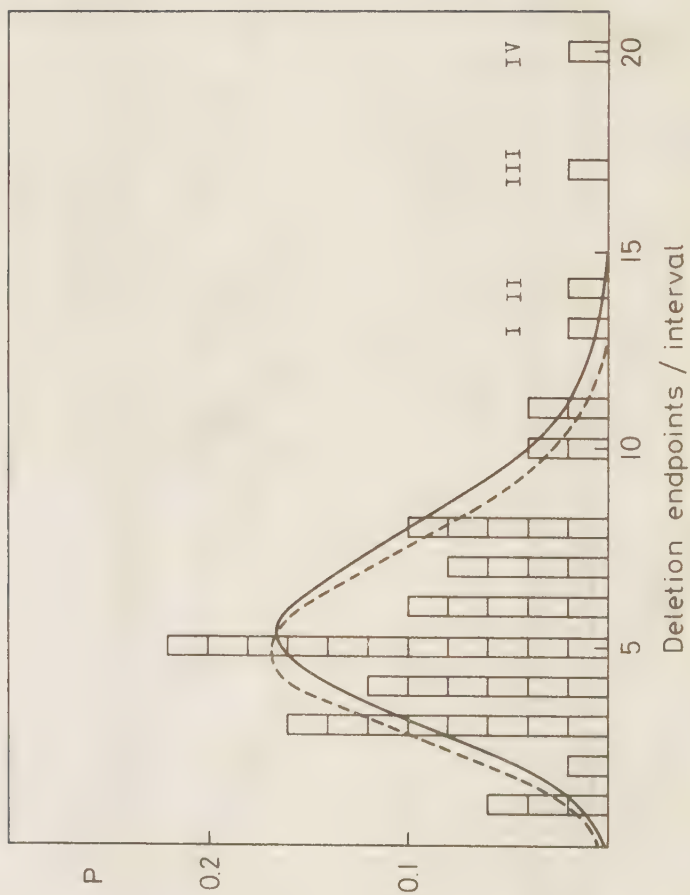
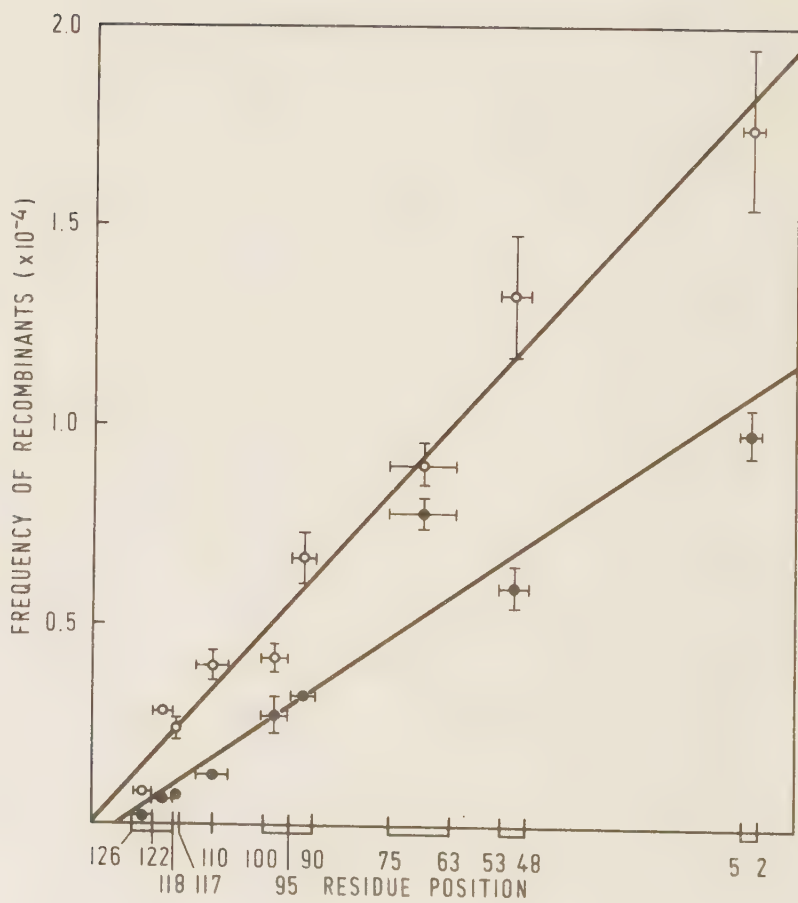


Figure 9

Figure
10

Figure
11



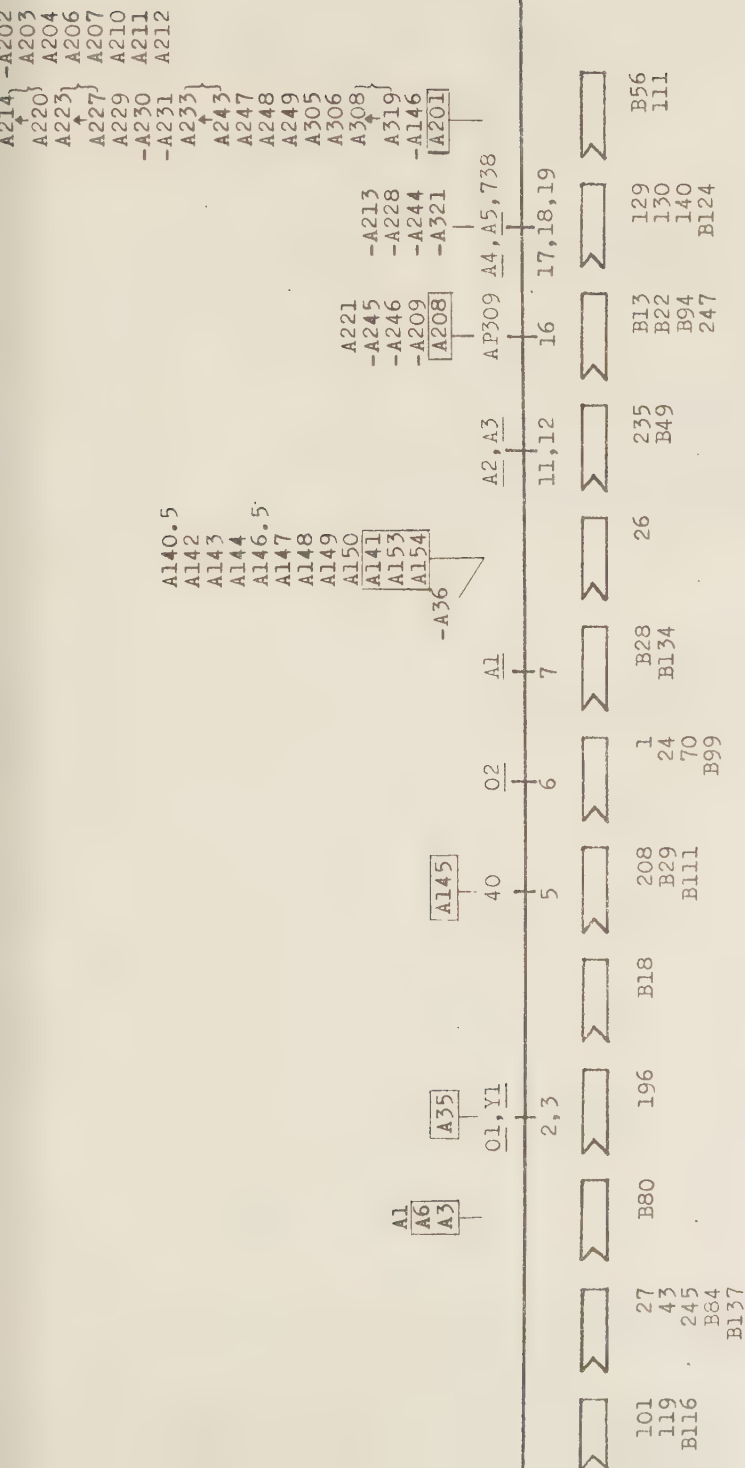


Figure 12a

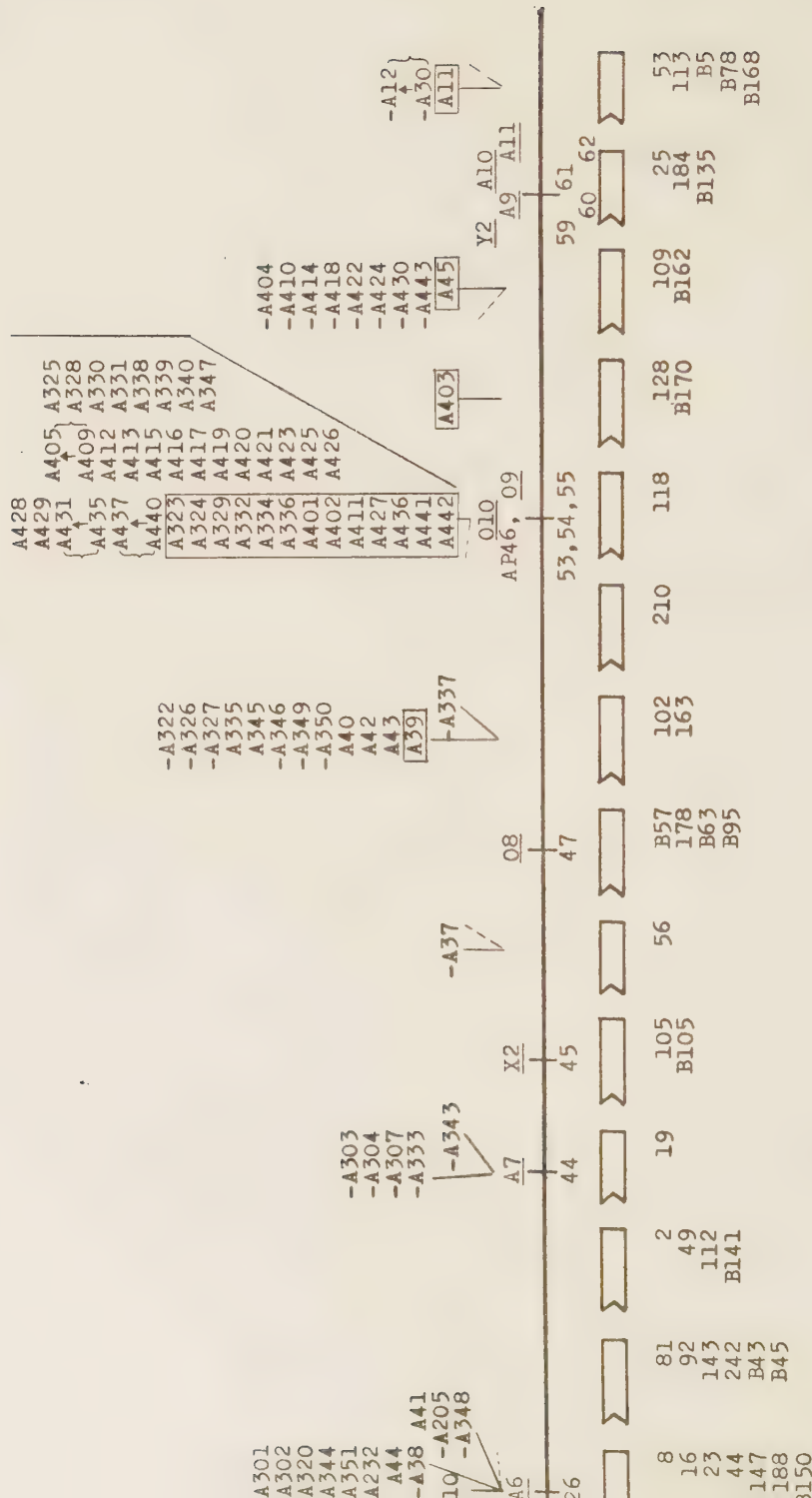


Figure 12b

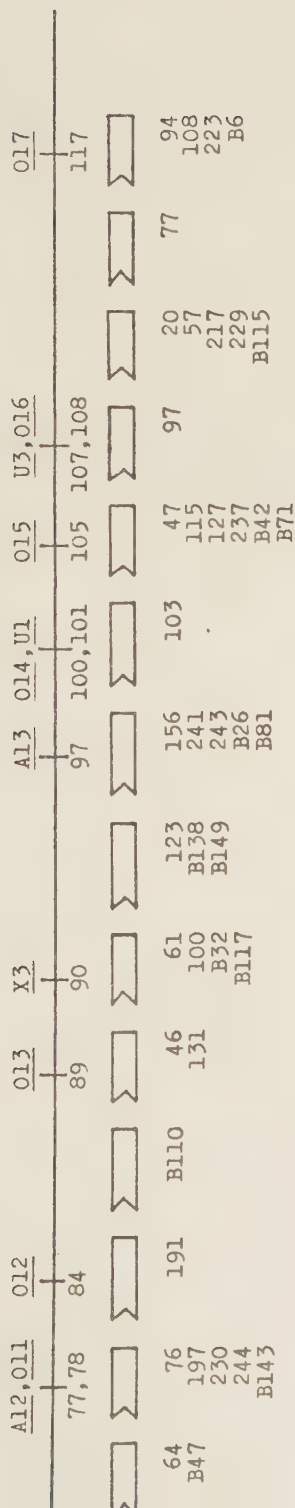


Figure 12c

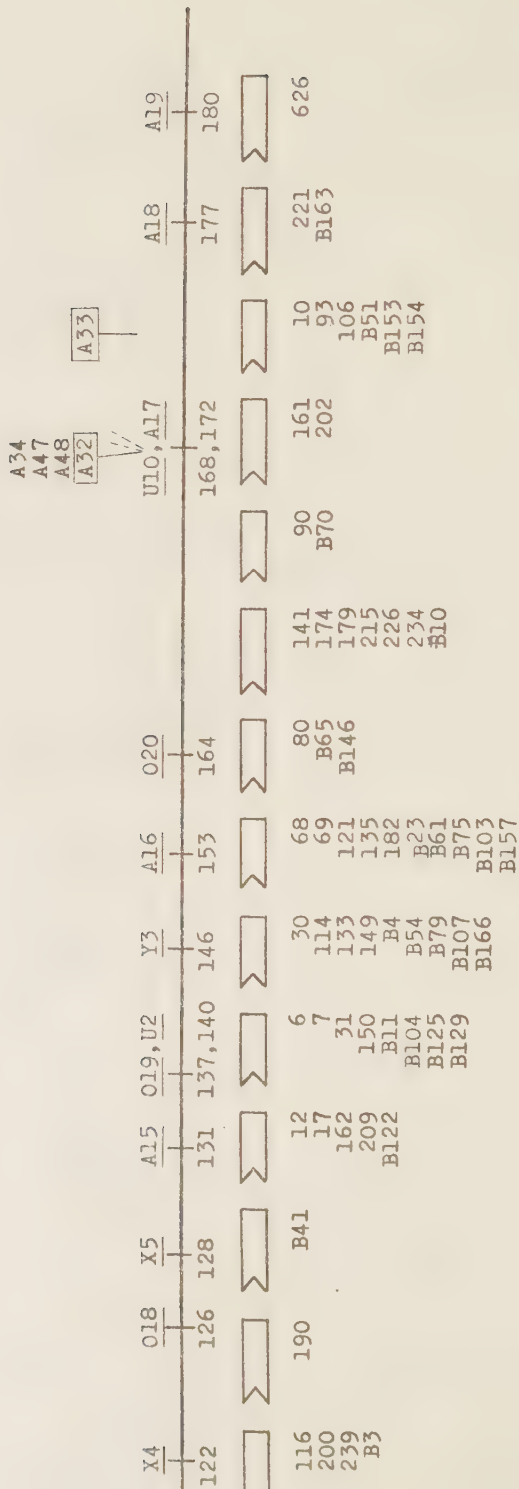
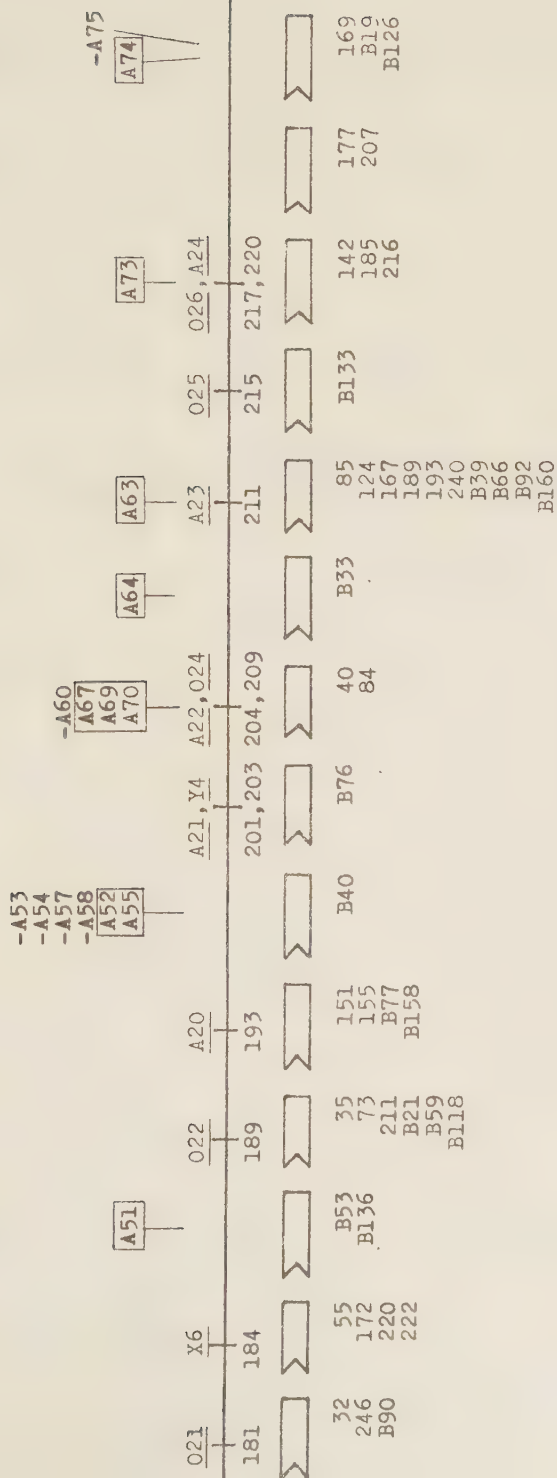


Figure 12d



-A77
-A78
-A79
-A80
-A81
-A82
-A83
-A84 }
-A89 }
-A91 }
-A94 }
-A100 }
-A103 }
-A106 }
-A107 }
-A176 }

A93 X7

A104

A92

A26

A25

A27

27,028 029

27,228 231

259

251

248

235

9
194
B14
B15
B20
B25
B91
B123

58
98
104
B82

50
63
95
132
B164

67

B1

250
B73
B128

139
B30
B64
B106
B151

238
B74

B36
B69

148
B113
B140

168
187

34
236
B72
B85
B89
B114
B142

144
165

75
78
83
87
107
153

Figure 12f

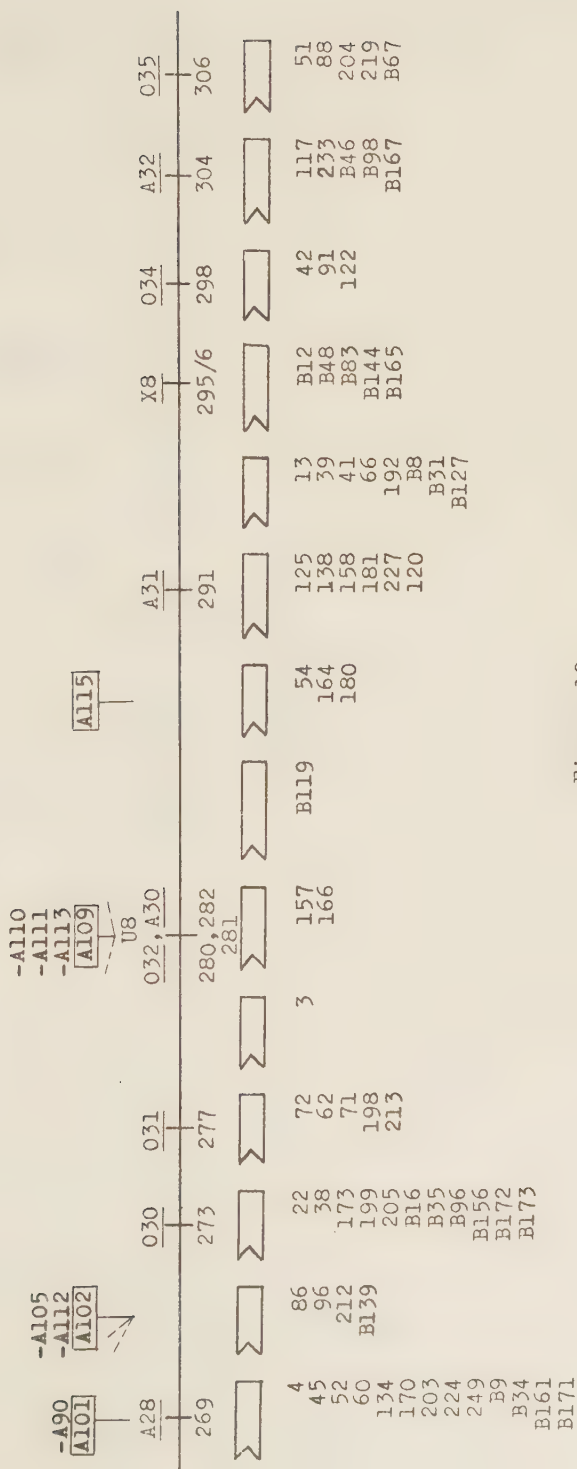


Figure 12g

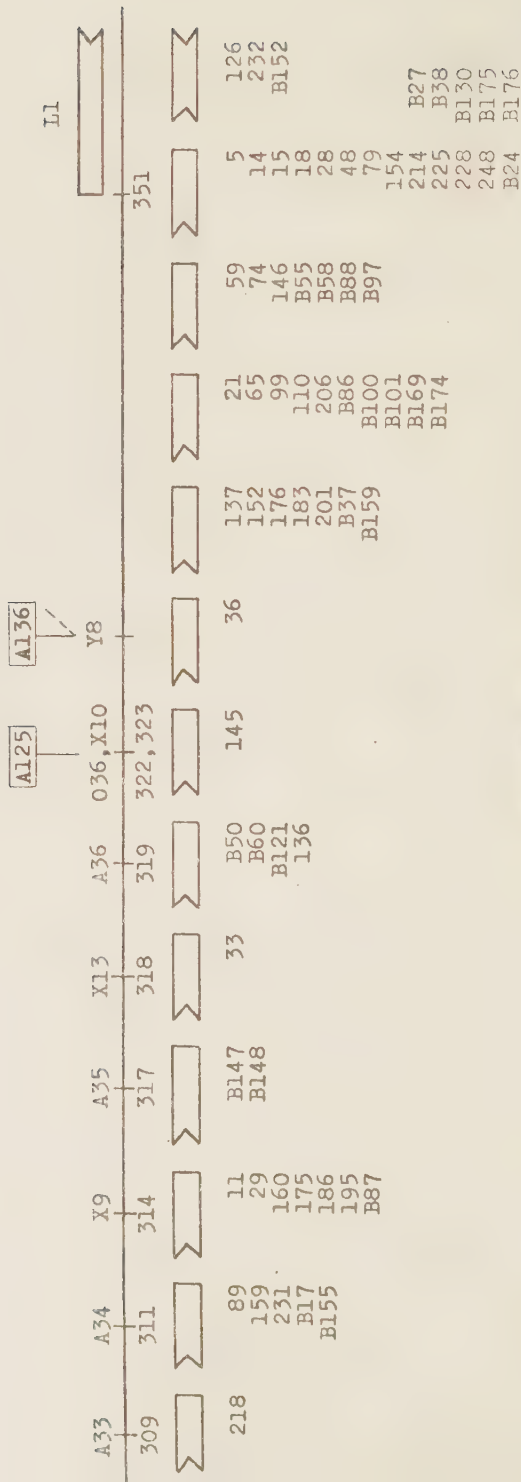


Figure 12h

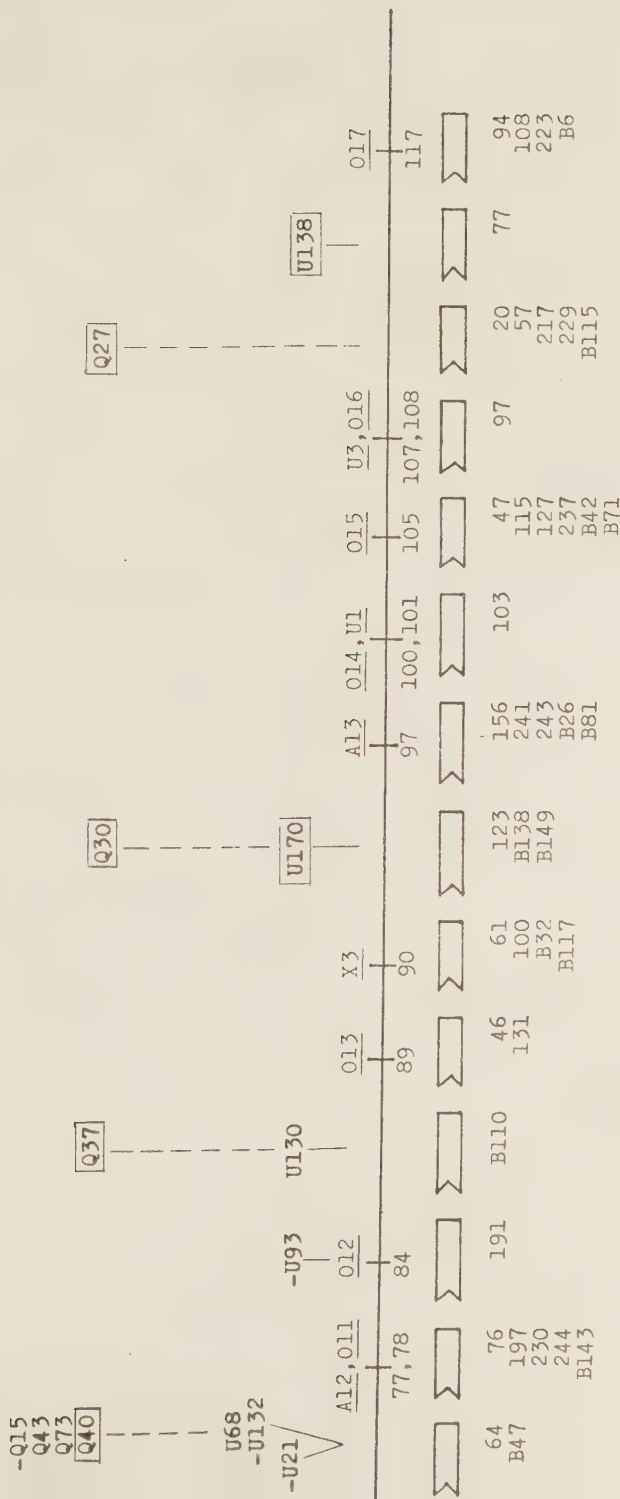


Figure 13c

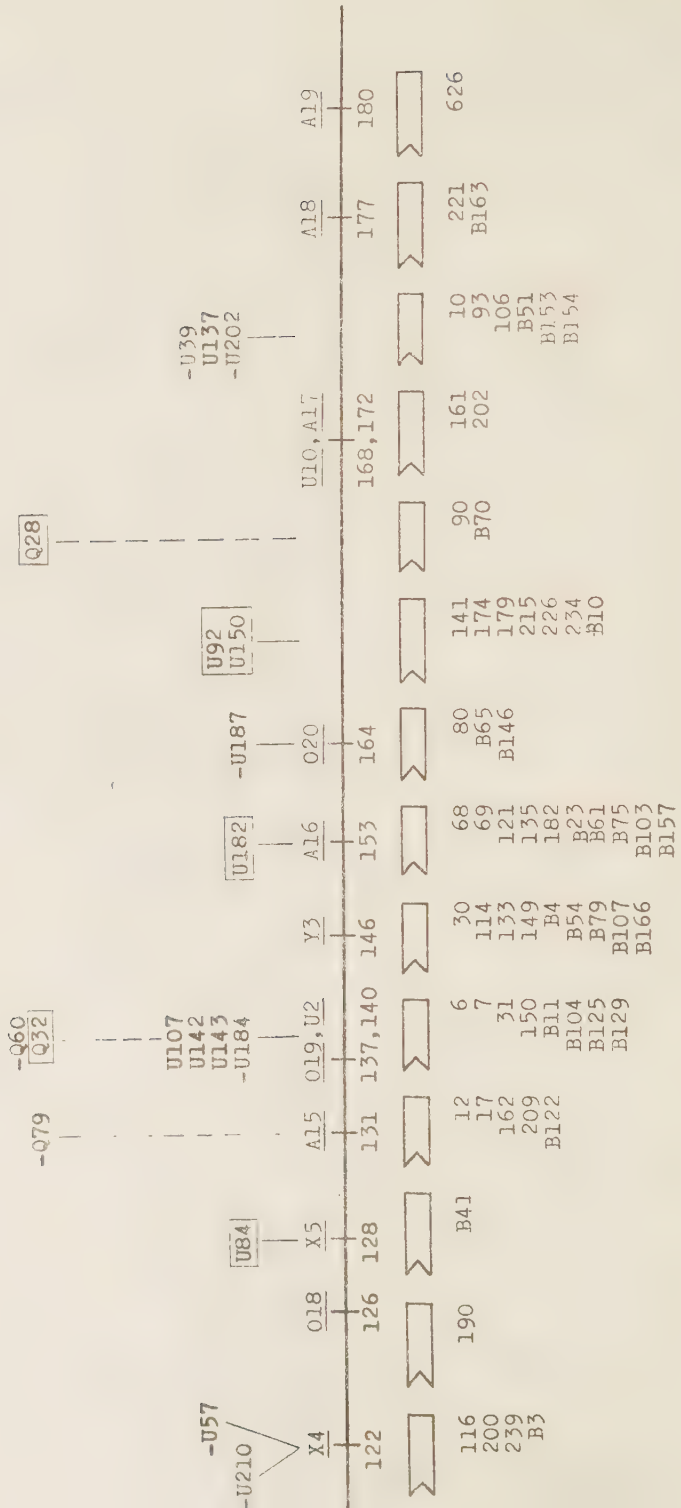


Figure 13d

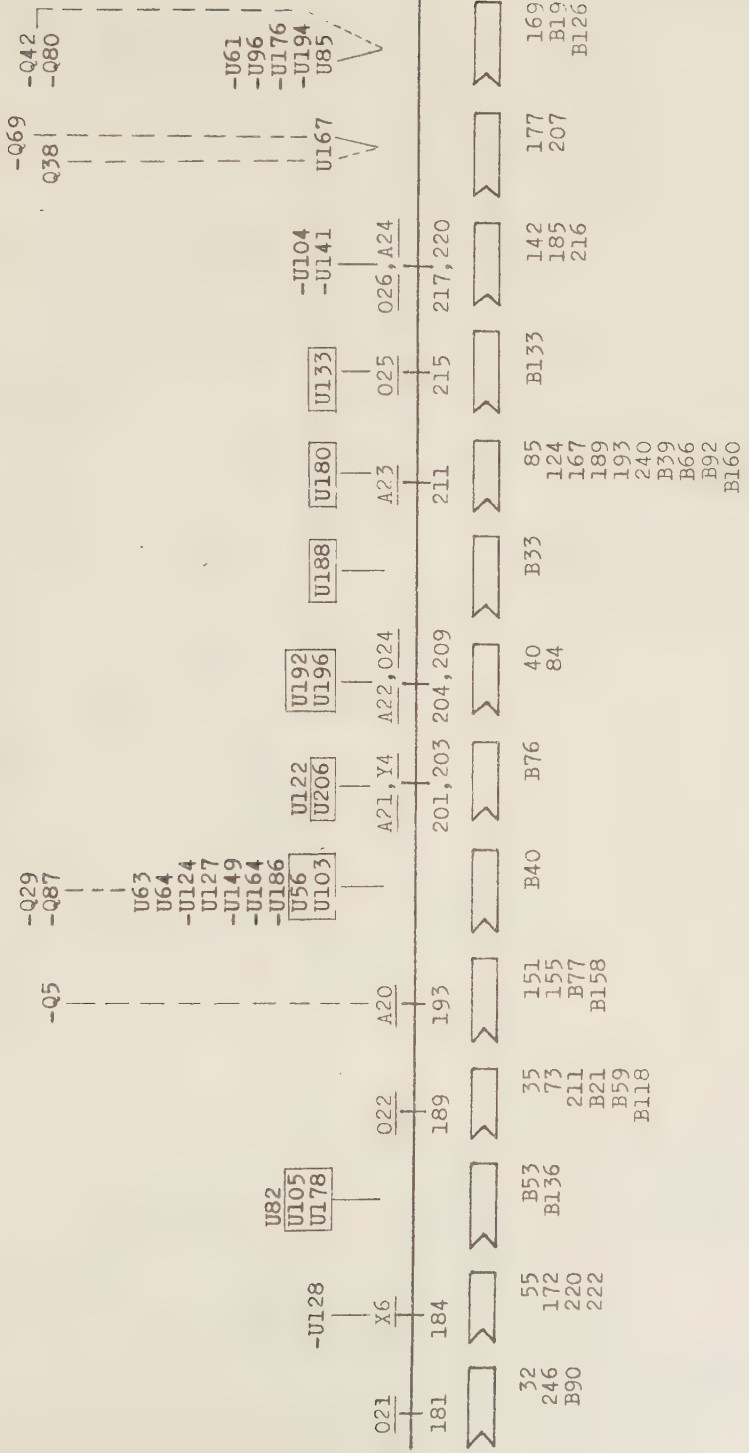


Figure 13e

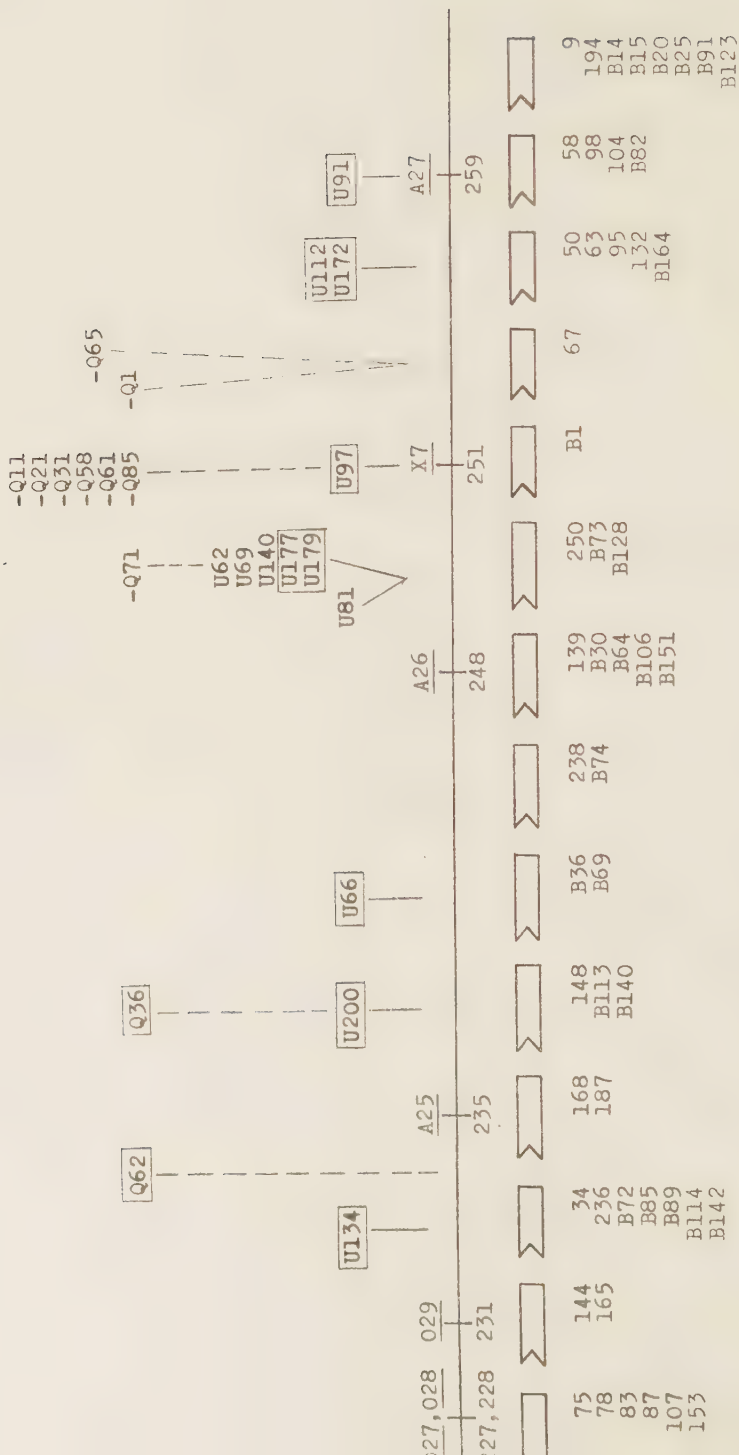


Figure 13f

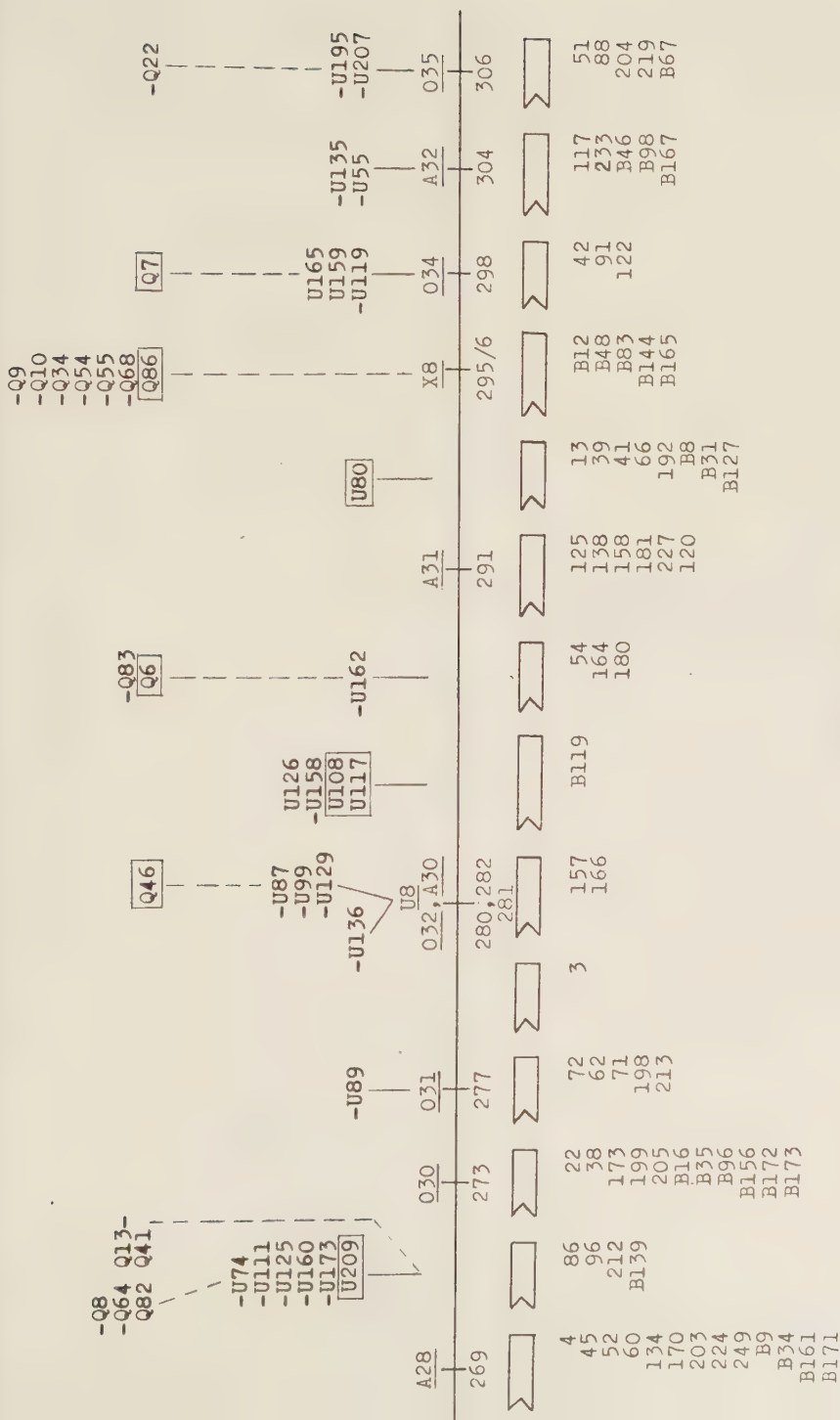
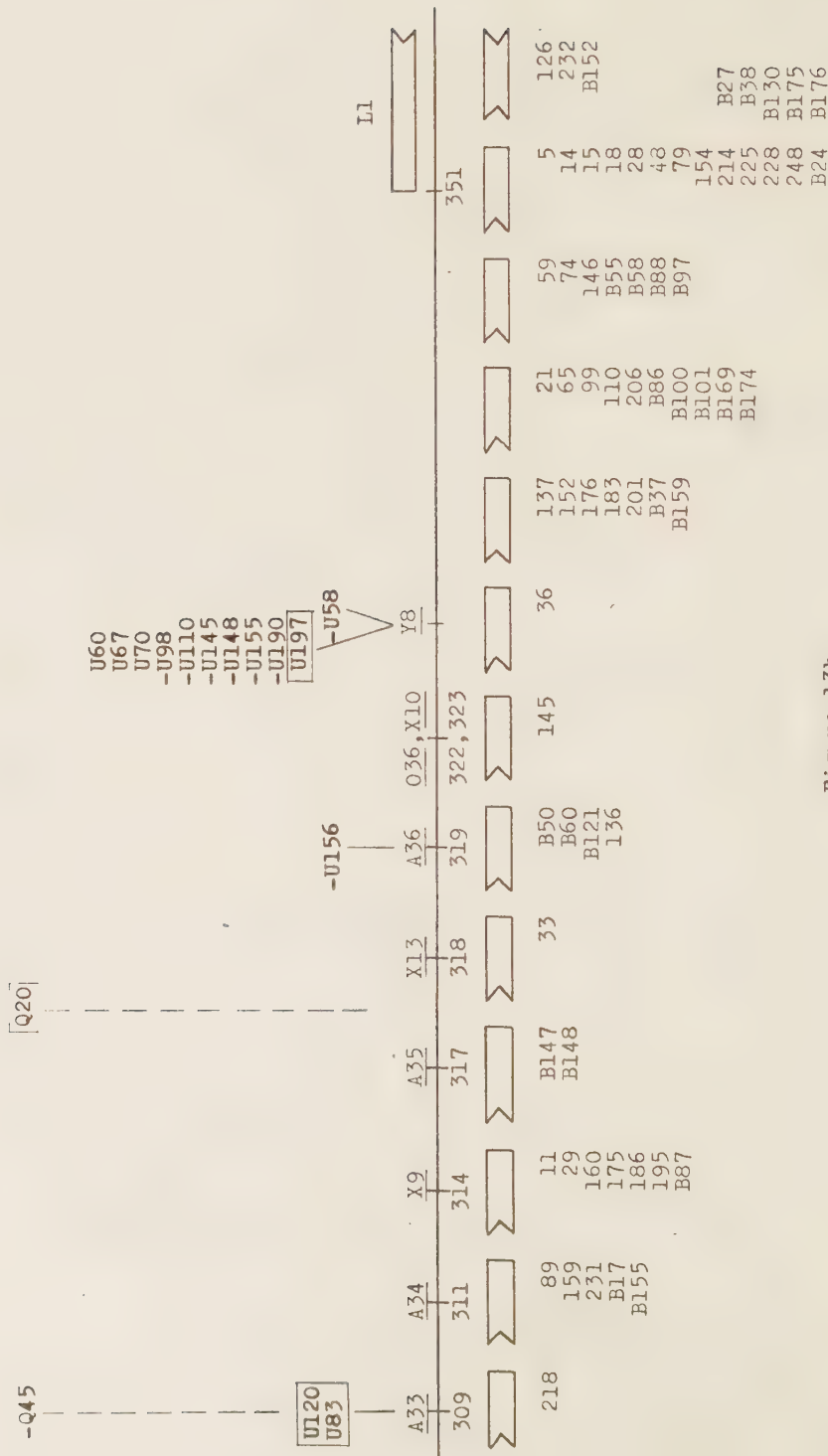


Figure 13g



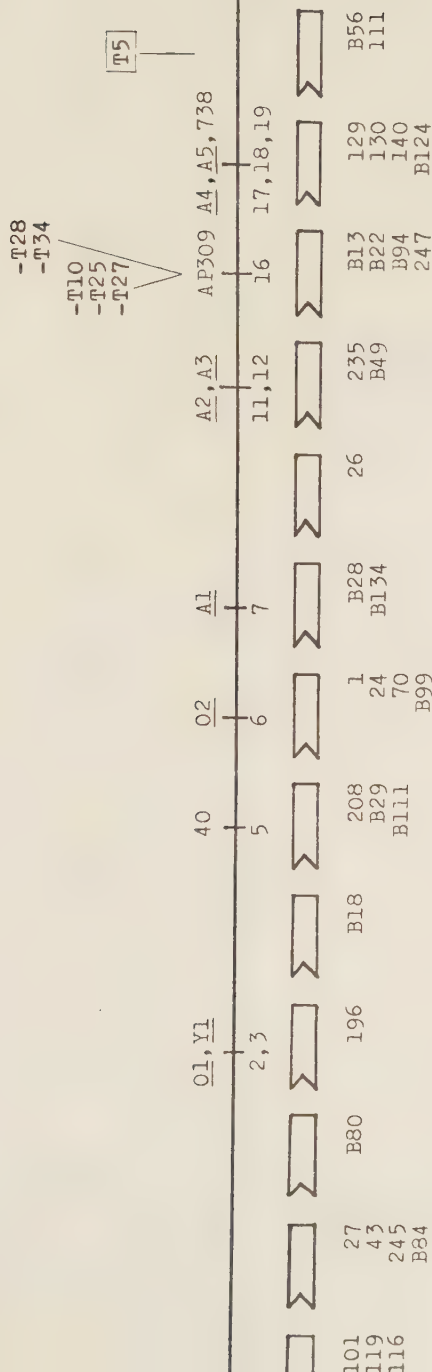


Figure 14a

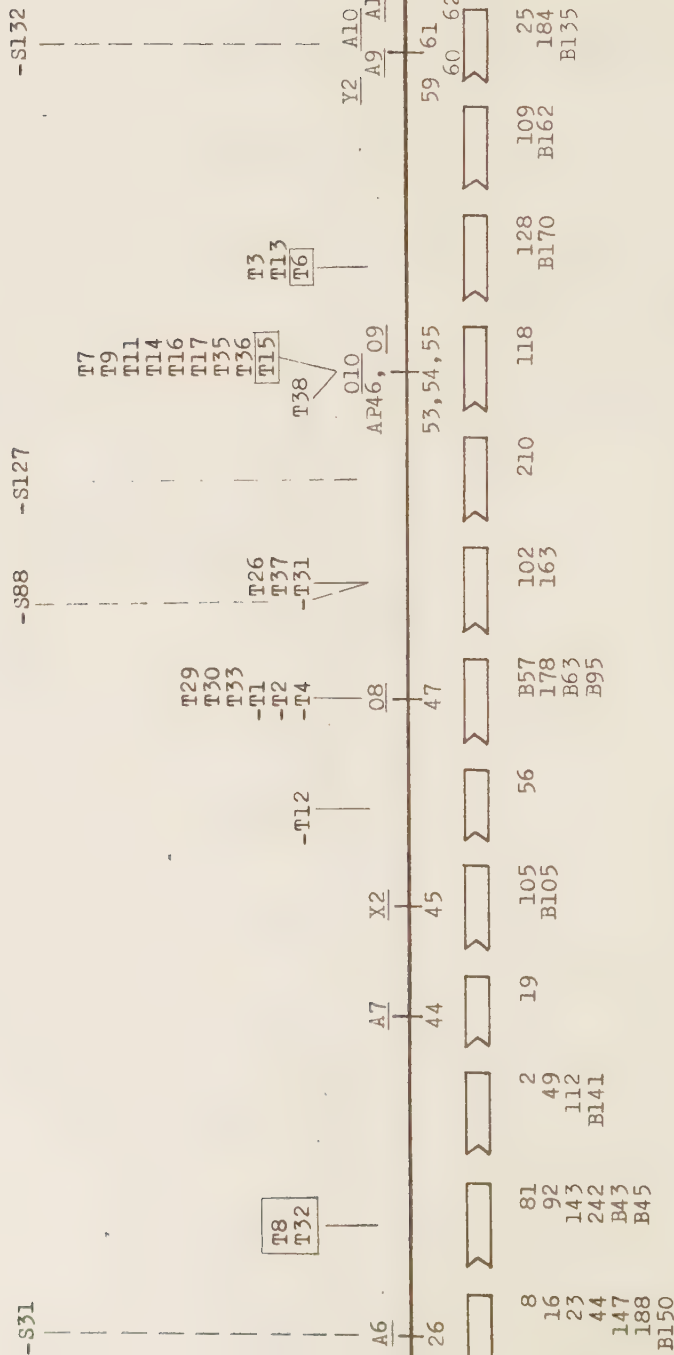


Figure 14b

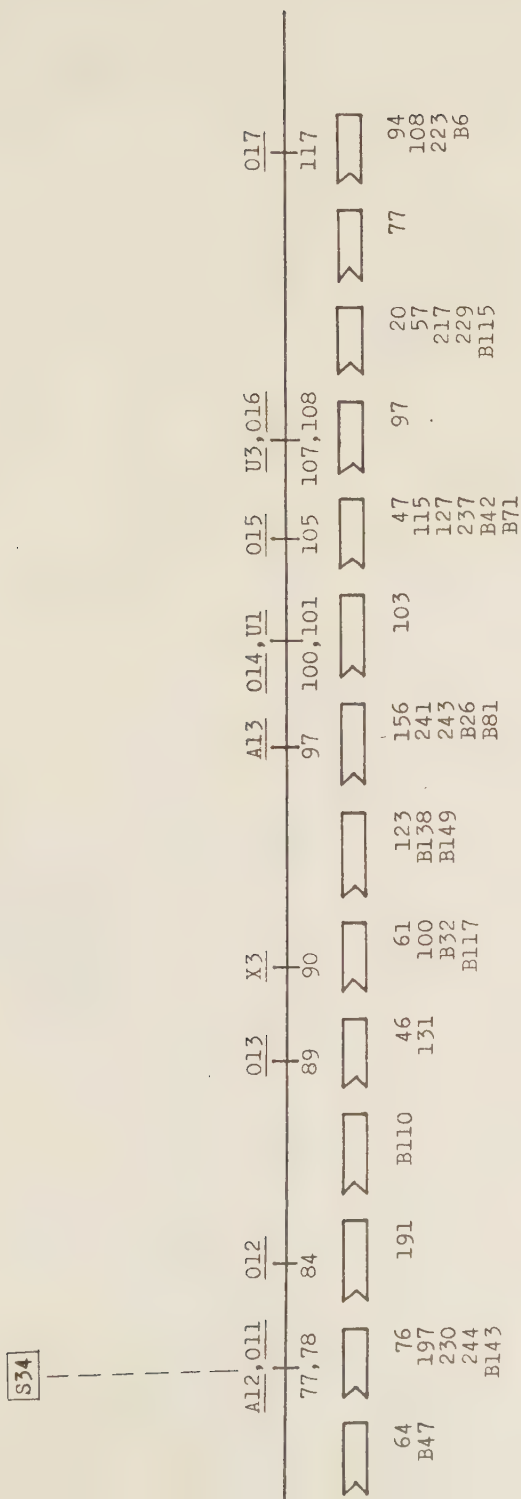


Figure 14c



Figure 14d

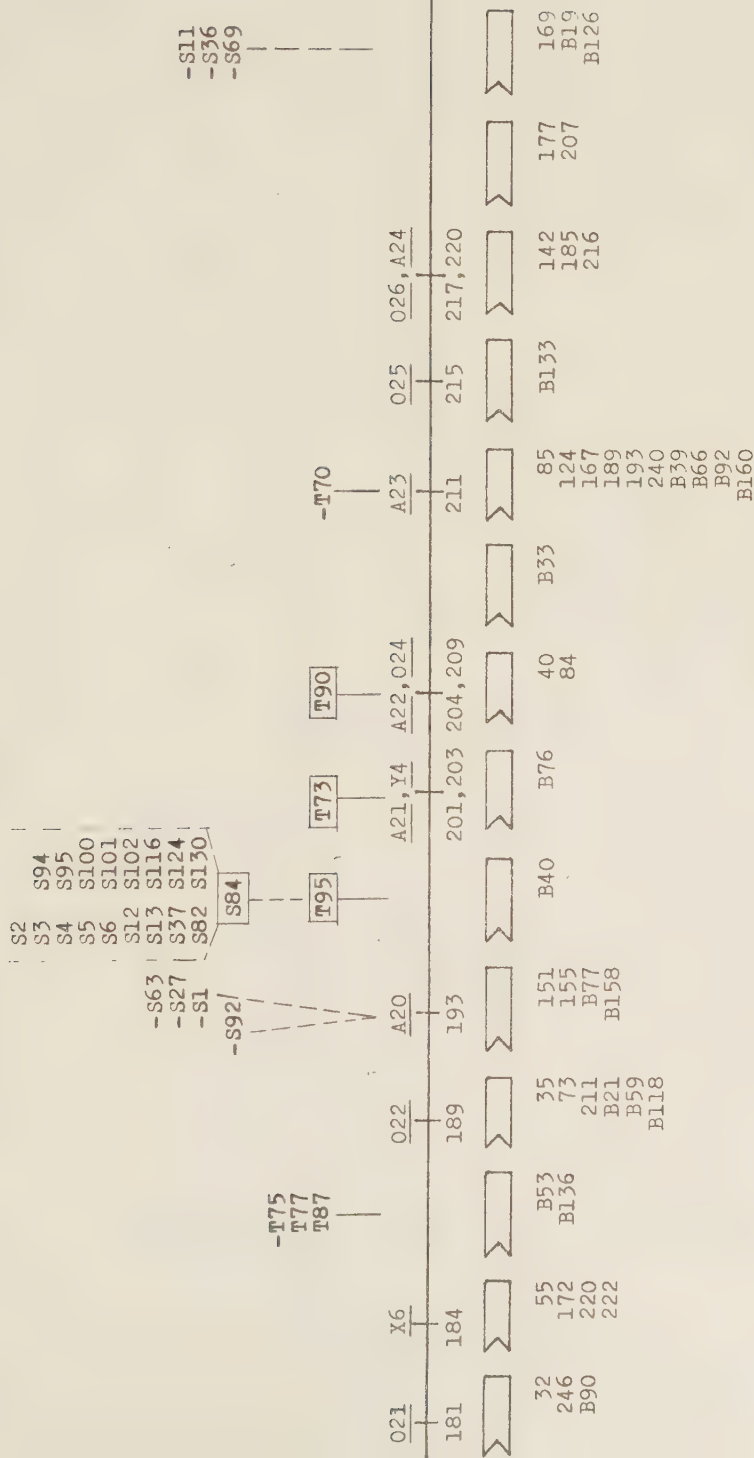


Figure 14e

[SZC]

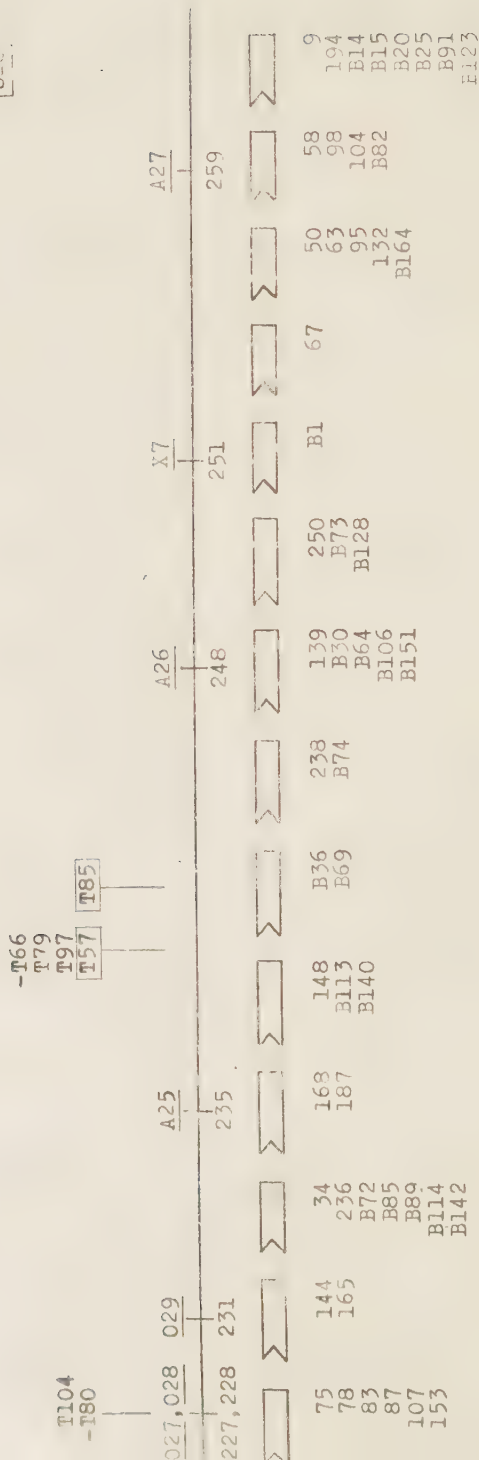


Figure 14

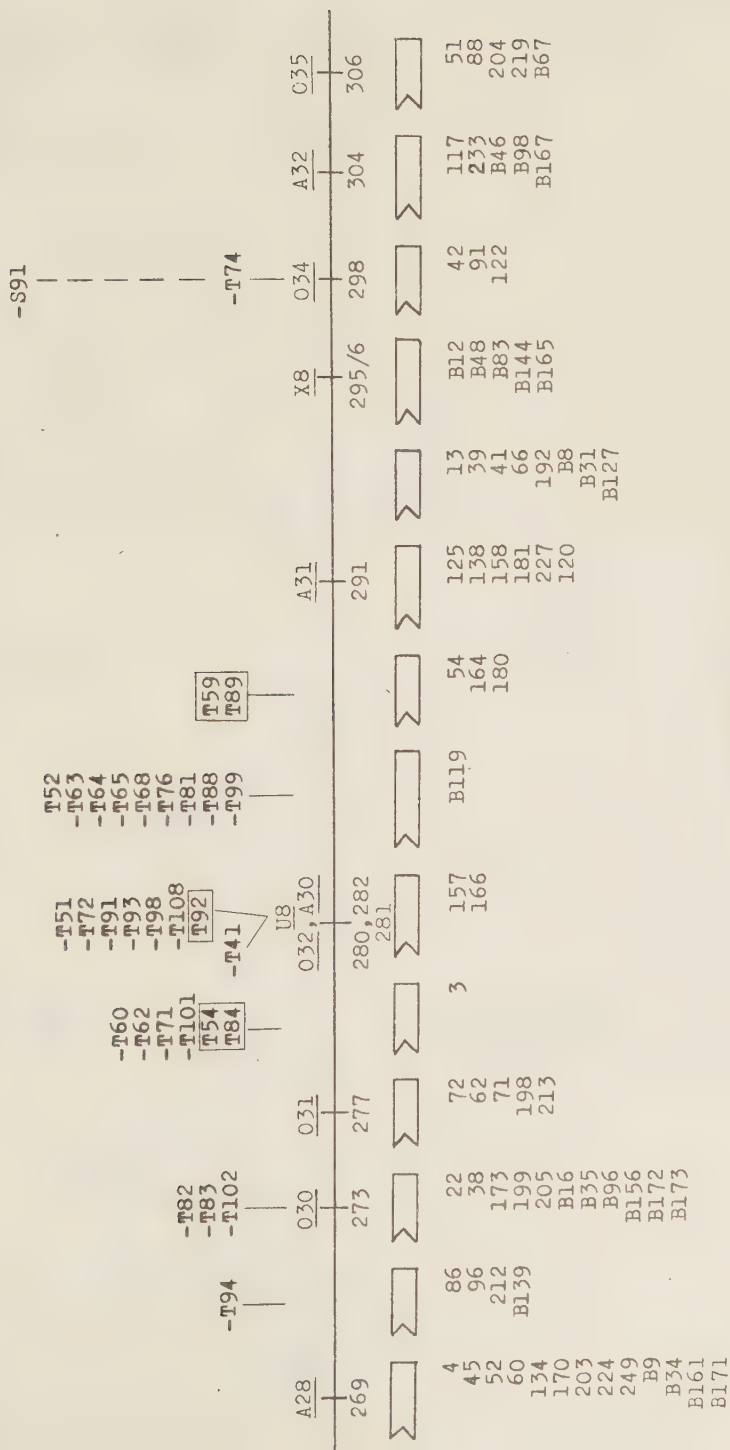


Figure 14g

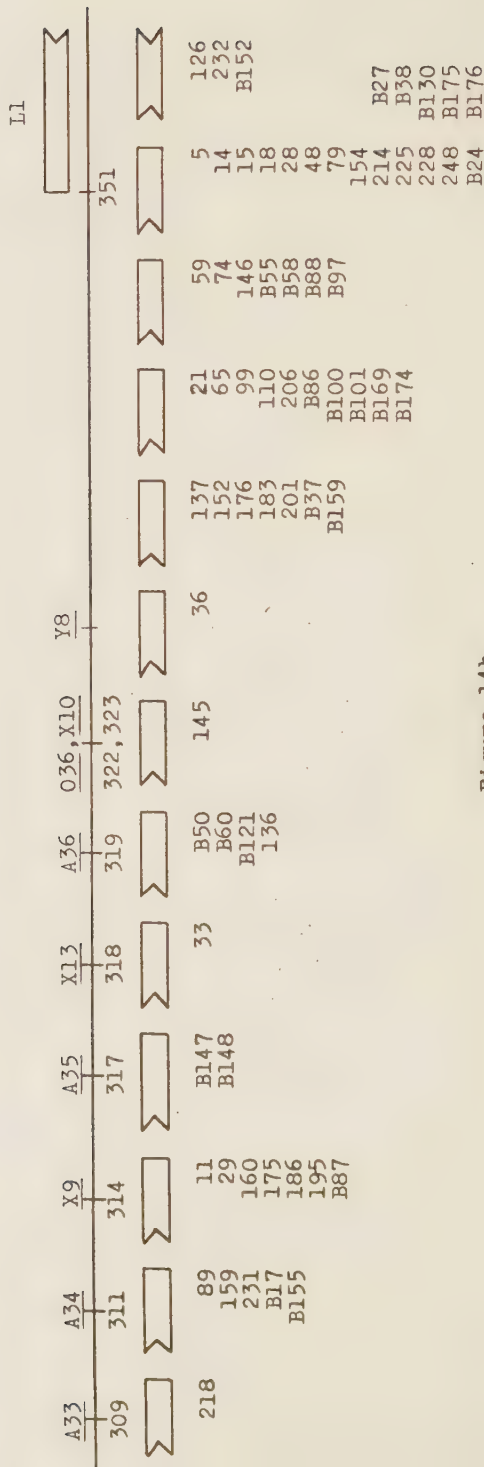
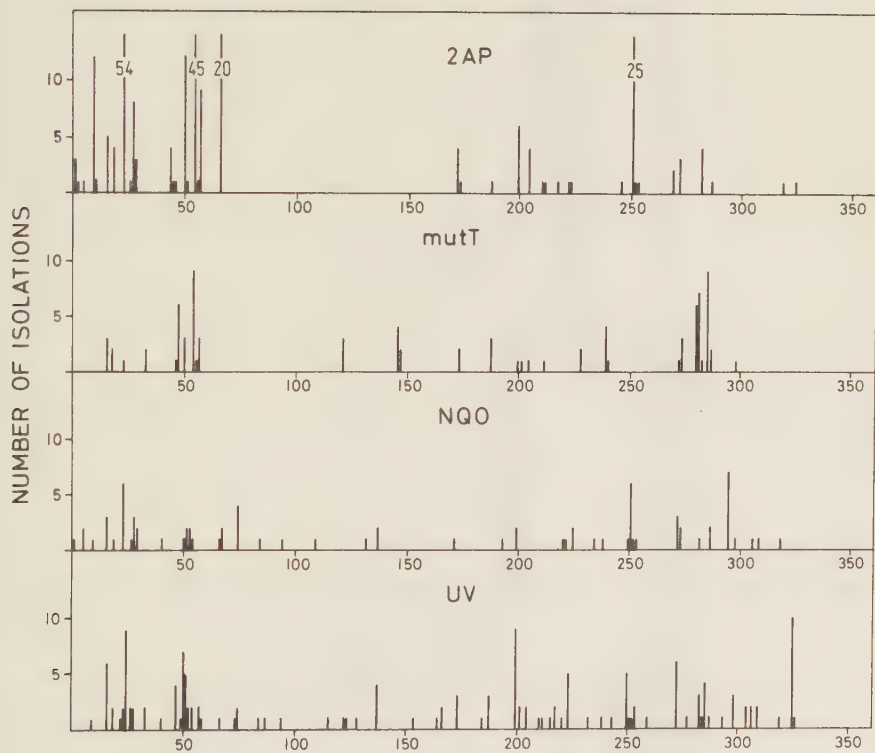


Figure 14a

Figure

15



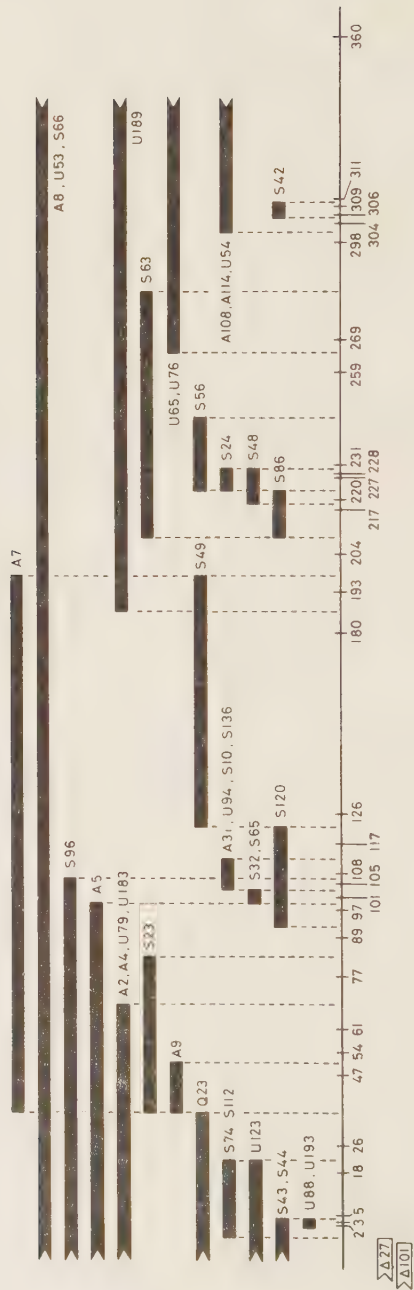


Figure
16

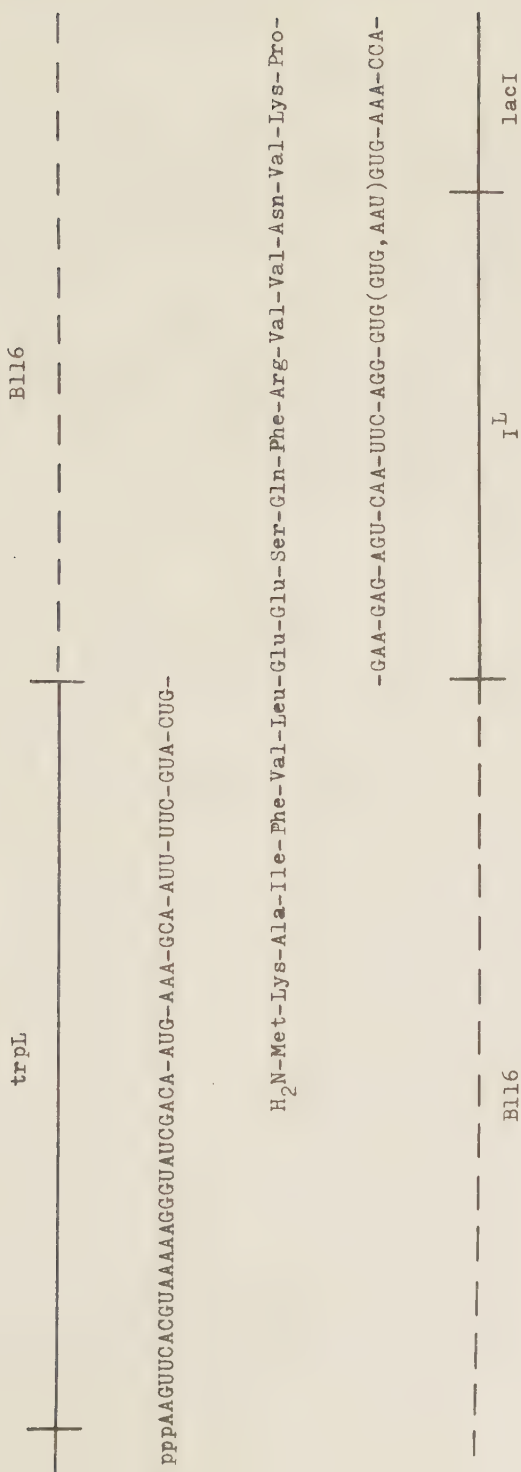


Figure 17

REFERENCES

- Adler K., Beyreuther K., Fanning E., Geisler N.,
Gronenborn B., Klemm A., Müller-Hill B., Pfahl M.,
Schmitz A. (1972). *Nature* 237, 322
- Allet B., Roberts R.J., Gesteland R.F., Solem R. (1974).
Nature 249, 217
- Ames B.N., Hartman P.E., Jacob F. (1963). *J. Mol. Biol.* 7, 23
- Anderson C.W. (1970). *Mutat. Res.* 9, 155
- Anderson W.B., Schneider A.B., Emmer M., Perlman R.L.,
Pastan I. (1971). *J. Biol. Chem.* 246, 5929
- Andrews K.J., Lin E.C.C. (1976). dans "Molecular Aspects of
Lac Operon Control", Cold Spring Harbor Lab.,
Abstracts, p. 54
- Arditti R.R., Scaife J.G., Beckwith J.R. (1968) *J. Mol.*
Biol. 38, 421
- Barkley M.D., Riggs A.D., Jobe A., Bourgeois S. (1975). *Biochem.* 14, 1700
- Beckendorf S.K., Wilson J.H. (1972). *Virology* 50, 315
- Beckwith J.R. (1964). *J. Mol. Biol.* 8, 427
- Beckwith J.R., Signer E.R. (1966). *J. Mol. Biol.* 19, 254
- Beckwith J.R., Signer E., Epstein W. (1966). Cold Spring
Harbor Symp. Quant. Biol. 23, 393
- Beckwith J.R. (1970). "The Lactose Operon", Cold Spring
Harbor Lab., New York, p. 5
- Beckwith J.R., Grodzicker T., Arditti R.R. (1972). *J. Mol.*
Biol. 69, 155
- Benzer S. (1961). *Proc. Natl. Acad. Sci. U.S.* 47, 403
- Bertrand K., Squires C., Yanofsky C. (1976). *J. Mol. Biol.*
103, 319
- Betz J.L., Sadler J.R. (1976) *J. Mol. Biol.* 105, 293
- Beyreuther K., Adler K., Geisler N., Klemm A. (1973). *Proc.*
Natl. Acad. Sci. U.S. 70, 3576
- Boos W., Schaedel P., Wallenfels K. (1967) *Eur. J. Biochem.*
1, 382
- Bourgeois S., Jobe A. (1970). "The Lactose Operon", Cold
Spring Harbor Lab., New York, p. 325
- Bourgeois S., Pfahl M. (1976). *Adv. Prot. Chem.* 30, 1
- Brenner S., Beckwith J.R., (1965). *J. Mol. Biol.* 13, 629
- Brenner S., Barnett L., Katz E.R., Crick F.H.C. (1967).
Nature 213, 449

- Campbell A., (1962). Adv. Genet. 11, 101
- Carlton B.C., Yanofsky C. (1963). J. Biol. Chem. 238, 2390
- Carlton B.C., Yanofsky C. (1965). J. Biol. Chem. 240, 690
- Chamness G.C., Willson C.D. (1970) J. Mol. Biol. 53, 561
- Chou P.Y., Adler A.J., Fasman G.D. (1975) J. Mol. Biol. 96, 29
- Cohn M. (1957). Bact. Rev. 21, 140
- Cohn M., Horibata K. (1959). J. Bact. 78, 624
- Coulondre C. Miller J.H. (1977a). J. Mol. Biol. 117, 525
- Coulondre C., Miller J.H. (1977b). J. Mol. Biol. 117, 576
- Coukell M.B., Yanofsky C. (1970). Nature 228, 633
- Coukell M.B., Yanofsky C. (1971). J. Bact. 105, 864
- Cox E.C., Yanofsky C. (1967). Proc. Natl. Acad. Sci. U.S. 58, 1895
- Craven G.R., Steers E., Anfinsen C.B. (1965). J. Biol. Chem. 240, 2468
- de Crombrughe B., Chen B., Anderson W.B., Gottesman M.E., Perlman R.L., Pastan I. (1971). J. Biol. Chem. 246, 7343
- Curtiss R. (1965). J. Bact. 89, 28
- Cuzin F., Jacob F. (1964). C. R. Acad. Sci. 258, 1350
- Davies J., Jacob F. (1968). J. Mol. Biol. 36, 413
- Demerec M., (1960). Proc. Natl. Acad. Sci. U.S. 46, 1075
- Dickson R.C., Abelson J. Barnes W.M., Reznikoff W.S. (1975). Science 187, 27
- Dickson R.C., Abelson J., Johnson (1977). J. Mol. Biol. 111, 65
- Drapeau G.R., Brammar W.J., Yanofsky C. (1968). J. Mol. Biol. 35, 357
- Emmer M., de Crombrughe B., Pastan I., Perlman R.L. (1970). Proc. Natl. Acad. Sci. U.S. 66, 480
- Fanning T.G. (1975) Biochem. 14, 2512
- Farabaugh P. (1978) en préparation
- Files J.G., Weber K., Miller J.H. (1974) Proc. Natl. Acad. Sci. U.S. 71, 667
- Files J.G., Weber K., Coulondre C., Miller J.H. (1975). J. Mol. Biol. 95, 327
- Fink G.R., Klopotoski T., Ames B.N. (1967). J. Mol. Biol. 30, 81
- Fowler A. (1972) J. Biol. Chem. 247, 5425

- Fukasawa T., Nikaido H. (1961) Biochem. Biophys. Acta 48, 470
- Fowler A.V., Zabin I. (1977). Proc. Natl. Acad. Sci. U.S. 74, 1507
- Fox C.F., Kennedy E.P. (1965). Proc. Natl. Acad. Sci. U.S. 54, 891
- Fox C.F., Carter J.R., Kennedy E.P. (1967). Proc. Natl. Acad. Sci. U.S. 57, 698
- Franklin N., Luria S. (1961). Virology 15, 299
- Franklin N. (1965). Biochem. Biophys. Res. Comm. 18, 910
- Franklin N. (1967). Genetics 55, 699
- Franklin N. (1971). dans "The Bacteriophage λ ", Cold Spring Harbor Laboratory, p. 175
- Ganem D. (1972). Honors Thesis, Harvard University
- Ganem D., Miller J.H., Files J.G., Platt T., Weber K. (1973). Proc. Natl. Acad. Sci. U.S. 70, 3165
- Gierer A. (1966). Nature 212, 1480
- Gilbert W., Müller-Hill B. (1966). Proc. Natl. Acad. Sci. U.S. 56, 1891
- Gilbert W., Müller-Hill B. (1967). Proc. Natl. Acad. Sci. U.S. 58, 2415
- Gilbert W., Müller-Hill B. (1970). "The Lactose Operon", Cold Spring Harbor Lab, p. 93
- Gilbert W., Maxam A. (1973). Proc. Natl. Acad. Sci. U.S. 70, 3581
- Gilbert W., Gralla J., Majors J., Maxam A. (1975). dans "Protein-Ligand Interactions", Walter de Gruyter, Berlin, p. 193
- Gilbert W. (1976a) dans "RNA Polymerase", éditeurs: Losick Chamberlin, Cold Spring Harbor Lab., New York, p. 197
- Gilbert W., Maxam A., Mirzabekov (1976b) dans "Control of Ribosome Synthesis", éditeurs: Kjielgaard et Maalge, The Alfred Benzon Symposium IX, Copenhagen, p. 139
- Gingery R., Echols H. (1968). Cold Spring Harbor Symp. Quant. Biol., 33, 721
- Goeddel D.C., Yansura D.G., Caruthers M.H. (1977). Proc. Natl. Acad. Sci. U.S. 74, 3292
- Gorini L. (1970). Ann. Rev. Genet. 4, 107
- Gottesman S., Beckwith J.R. (1969). J. Mol. Biol. 44, 117
- Gottesman E.M., Weisberg R.A. (1971). "The Bacteriophage λ ", Cold Spring Harbor Laboratory, p. 113

- Gratia J.P. (1964). Ann. Inst. Pasteur 107, 132
- Guest J.R., Yanofsky C. (1965). J. Biol. Chem. 240, 679
- Gursky A.V., Tumanyan V.G., Zasedatelev A.S., Zhuze A.L.,
Grokhovsky S.L., Gottikh B.P. (1976). Mol. Biol.
Reports 2, 413
- Hartman P.E., Hartman Z., Stahl R.C. (1971) Adv. Genet, 16, 1
- von Hippel P., Revzin A., Gross C., Wang A.C. (1975). dans
"Protein-Ligand Interactions" W. de Gruyter,
Berlin, p. 270
- Hopkins J.D. (1974) J. Mol. Biol. 87, 715
- Imamoto F., Ito J., Yanofsky C. (1966). Cold Spring Harbor
Symp. Quant. Biol., 31, 235
- Inselburg J., (1967). J. Bact. 94, 1266
- Ippen K., Miller J.H., Scaife J. Beckwith J.R. (1968).
Nature 217, 825
- Itikawa H., Demerec M. (1967). Genetics 55, 63
- Jacob F., Monod J. (1961). J. Mol. Biol. 3, 318
- Jacob F., Ullmann A., Monod J. (1964). C.R. Acad. Sci. 258,
3125
- Jacob F., Ullmann A., Monod J. (1965). J. Mol. Biol. 13, 704
- Jobe A., Riggs A.D., Bourgeois S. (1972). J. Mol. Biol.
64, 181
- Jobe A., Bourgeois S. (1972a). J. Mol. Biol. 69, 397
- Jobe A., Bourgeois S. (1972b) J. Mol. Biol. 72, 139 X86
- Jobe A., Bourgeois S. (1973). J. Mol. Biol. 75, 303
- Jobe A., Sadler J.R., Bourgeois S. (1974) J. Mol. Biol. 85,
231
- Jones O.W., Berg P. (1966). J. Mol. Biol. 22, 199
- Jones T.D.H., Kennedy E.P. (1969). J. Biol. Chem. 244, 5981
- Kao-Huang Y., Revzin A., Butler A.P., O'Conner P., Noble
D.W., v. Hippel P., (1977) Proc. Natl. Acad. Sci.
U.S., 74, 4228
- Kennedy E.P. (1970). dans "The Lactose Operon", Cold Spring
Harbor Laboratory, p. 49
- Kennell D., Riezman H. (1977). J. Mol. Biol. 114, 1
- Kepes A. (1963). Biochim. Biophys. Acta 76, 293

- Kleckner N. (1977) Cell, 11, 11
- Korn L.J., Yanofsky C. (1976). J. Mol. Biol. 103, 395
- Lam S.T., Stahl M.M., McMillin K.D., Stahl F.W. (1974)
Genetics 77, 425
- Lee F., Yanofsky C. (1977). Proc. Natl. Acad. Sci. U.S.
74, 4365
- Lillis M., (1977). J. Mol. Biol. 117, 568
- Lin S.Y., Riggs A.D. (1972). J. Mol. Biol. 72, 671
- Lin S.Y., Riggs A.D. (1975a). Biochem. Biophys. Res. Comm.
62, 704
- Lin S.Y., Riggs A.D. (1975b). Cell. 4, 107
- Magasanik B., (1961). Cold Spring Harbor Symp. Quant. Biol.
26, 249
- Makman R.S., Sutherland E.W. (1965). J. Biol. Chem. 240, 1309
- Malamy M., (1966). Cold Spring Harbor Symp. Quant. Biol.
31, 189
- Maurer R., Maniatis T., Ptashne M. (1974). Nature 249, 221
- Meyer B., Kleid D., Ptashne M., (1975). Proc. Natl. Acad.
Sci. U.S. 72, 4785
- Miller J.H., Ippen K., Scaife J., Beckwith J. (1968a).
J. Mol. Biol. 38, 413
- Miller J.H., Beckwith J.R., Müller-Hill B. (1968b). Nature
220, 1287
- Miller J.H. (1970). The Lactose Operon, Cold Spring Harbor
Lab., p. 133
- Miller J.H., Reznikoff W.S., Silverstone A., Ippen K.,
Signer E.R., Beckwith J.R. (1970). J. Bact. 104, 1273
- Miller J.H. (1972). "Experiments in Molecular Genetics"
Cold Spring Harbor Laboratory, New York
- Miller J.H., Ganem D., Lu P., Schmitz A. (1977). J. Mol.
Biol. 109, 275
- Mirzabekov A.D., Kolchinsky A.M. (1974). Mol. Biol. Rep.
1, 379
- Mitchell D.H., Reznikoff W.S., Beckwith J.R. (1975).
J. Mol. Biol. 93, 331
- Mitchell D.H., Reznikoff W.S., Beckwith J.R. (1976).
J. Mol. Biol. 101, 441

- Monod J., Cohen-Bazire G., Cohn M. (1951). *Biochim. Biophys. Acta*, 7, 585
- Moore C.W., Sherman F. (1975). *Genetics* 79, 397
- Moore C.W., Sherman F. (1977). *Genetics* 85, 1
- Müller-Hill B., Rickenberg H.W., Wallenfels K. (1964). *J. Mol. Biol.* 10, 303
- Müller-Hill B., Crapo, Gilbert W. (1968). *Proc. Natl. Acad. Sci. U.S.* 59, 1259
- Müller-Hill B., Beyreuther K., Gilbert W. (1971). *Methods in Enzymology*, vol. 21D p. 483
- Müller-Hill B., Fanning T., Geisler N., Gho D., Kania J., Kathmann P., Meissner H., Schlotmann M., Schmitz A., Triesch I., Beyreuther K. (1975). dans "Protein-Ligand Interactions", Walter de Gruyter, Berlin p. 211
- Müller-Hill B., (1975). *Progr. Biophys. Mol. Biol.*, Pergamon Press England vol. 30, p. 227
- Newton A., Beckwith J.R., Zpser D., Brenner S. (1965). *J. Mol. Biol.* 14, 290
- Norkin L.C. (1970). *J. Mol. Biol.* 51, 633
- Oshima Y., Horiuchi T., Yamagida M., (1975). *J. Mol. Biol.* 91, 515
- Osborn M., Person S., Phillips S., Funk F. (1967). *J. Mol. Biol.* 26, 437
- Pardee A.B., Jacob F., Monod J. (1959) *J. Mol. Biol.* 1, 165
- Parkinson J.S., Huskey R.J. (1971). *J. Mol. Biol.* 56, 369
- Perlman R.L., Pastan I. (1969). *Biochem. Biophys. Res. Comm.* 37, 151
- Pfahl M. (1972). *Genetics* 72, 393
- Pfahl M., Stockter C., Gronenborn B. (1974). *Genetics* 76, 669
- Pfahl M. (1976). *J. Mol. Biol.* 106, 857
- Pfeifer D., Oellermann R. (1967). *Mol. Gen. Genet.* 99, 248
- Pfeifer D., Rosypal S. (1969). *Mol. Gen. Genet.* 104, 116
- Pfeifer D., Hirsch H.J., Bergmann D., Hamlaoui M. (1974). *Mol. Gen. Genet.* 132, 203
- Platt T., Weber K., Ganem D., Miller J.H. (1972). *Proc. Natl. Acad. Sci. U.S.* 69, 897

- Platt T., Weber K., Ganem D., Miller J.H..(1973). J. Biol. Chem. 248, 110
- Platt T., Squires C., Yanofsky C. (1976). J. Mol. Biol. 103, 411
- Prakash L., Stewart J.W., Sherman F. (1974). J. Mol.Biol. 85,51
- Pribnow D. (1975). J. Mol. Biol. 92, 419
- Reif H.J., Saedler H. (1975). Mol. Gen. Genet. 137, 17
- Reif H.J., Saedler H. (1977). dans "DNA Insertion Elements, Plasmids and Episomes", Cold Spring Harbor Lab., New York, p.81
- Reznikoff W.S., Miller J.H., Scaife J.G., Beckwith J.R. (1969) J. Mol. Biol. 43, 201
- Richter P.H., Eigen M. (1974). Biophys. Chem. 2, 255
- Riggs A.D., Bourgeois S. (1968). J. Mol. Biol. 24, 361
- Riggs A.D., Newby R.F., Bourgeois S., Cohn M. (1968). J. Mol. Biol. 24, 365
- Riggs A.D., Suzuki H., Bourgeois S. (1970a). J.Mol.Biol. 48,67
- Riggs A.D., Bourgeois S., Cohn M. (1970b). J. Mol. Biol. 53,401
- Riggs A.D., Newby R.F., Bourgeois S. (1970c). J. Mol. Biol. 51, 303
- Riggs A.D., Reiness G., Zubay G. (1971). Proc. Natl. Acad. Sci. U.S. 68, 1222
- Riggs A.D., Lin S., Wells . (1972). Proc. Natl. Acad. Sci. U.S. 69, 761
- Ronen A., Salts Y. (1971). Virology 45, 496
- Rosenberg J.M., Khallai O.B., Kopka M.L., Dickerson R.E., Riggs A.D. (1977). Nucl. Ac. Res. 4, 567
- Sadler J.R., Smith T.F. (1971). J. Mol. Biol. 62, 139
- Sarabhai A.S., Stretton A.O., Brenner S., Bolle A. (1964). Nature 201, 13
- Scaife J., Beckwith J.R. (1966).. Cold Spring Harbor Symp. Quant. Biol. 31, 403
- Schaller H., Gray C., Herrmann K. (1975). Proc. Natl. Acad. Sci. U.S. 72, 737
- Schleif R. (1972). Proc. Natl. Acad. Sci. U.S. 69, 3479
- Schmeissner U., Ganem D., Miller J.H. (1977a). J. Mol. Biol. 109, 303
- Schmeissner U., Ganem D., Miller J.H. (1977b). J. Mol. Biol. 117, 572

- Schmitz A., Schmeissner U., Miller J.H., Lu P. (1976). J. Biol. Chem. 251, 3359
- Schmitz A., Coulondre C., Miller J.H., (1978). J. Mol. Biol. sous presse
- Schwartz D., Beckwith J.R. (1969). Genetics 61, 371
- Schwartz D., Beckwith J.R. (1970). dans "The Lactose Operon", Cold Spring Harbor Lab., p. 417
- Shapiro J.A., Adhya S.L. (1969). Genetics 62, 249
- Sherman F., Jackson M., Liebman S.W., Schweingruber A.M., Stewart J.W. (1972). Genetics 81, 51
- Shimada K., Weisberg R.A., Gottesman M.E. (1972). J. Mol. Biol. 63, 483
- Signer E.R., Beckwith J.R., Brenner S. (1965). J. Mol. Biol. 14, 153
- Smith T.F., Sadler J.R. (1971). J. Mol. Biol. 59, 273
- Sommer H. (1977). J. Mol. Biol. 109, 299
- Sommer H., Schmitz A., Schmeissner U., Miller J.H., Wittmann H.G. (1978). J. Mol. Biol. sous presse
- Spudich J.A., Horn V., Yanofsky C. (1976). J. Mol. Biol. 53, 49
- Squires C., Lee F., Bertrand K., Squires C.L., Bronson M.J., Yanofsky C. (1976). J. Mol. Biol. 103, 351
- Stadler D. (1973). Ann. Rev. Genet. 7, 113
- Stadler D., Karija B. (1973). Genetics 75, 423
- Steege D. (1977). Proc. Natl. Acad. Sci. U.S. 74, 4163
- Steitz T.A., Richmond T.J., Wise D., Engelman D. (1977). Proc. Natl. Acad. Sci. U.S. 71, 593
- Sternberg M.J.F., Thornton J.M. (1978). Nature 271, 15
- Tessman I. (1962). J. Mol. Biol. 5, 442
- Tessman I. (1965). Genetics 51, 63
- Villarejo M. (1976). dans "Molecular Aspects of lac Operon Control", Cold Spring Harbor Lab., Abstracts, p. 55
- Wallenfels K., Malhotra O.P. (1961). Adv. Carbohydr. Chem., Acad. Press, New York, p. 239
- Wang J.C., Barkley M.D., Bourgeois S. (1974). Nature 251, 247
- Wang A.C., Revzin A., Butler A.P., von Hippel P. (1977). Nucl. Ac. Res. 4, 1579
- Weber K., Platt T., Ganem D., Miller J.H. (1972). Proc. Natl. Acad. Sci. U.S. 69, 3624

- Weber K., Piles J.G., Platt L., Ganem D., Miller J.H. (1975).
dans "Protein-Ligand Interactions", Walter de Gruyter,
Berlin, p. 228
- Weisberg R.A., Gottesman M.E. (1969). J. Mol. Biol. 46, 565
- Willson C.D., Perrin D., Cohn M., Jacob F., Monod J. (1964).
J. Mol. Biol. 8, 582
- Wilson F.H., Kashket E.R. (1969). Biochim. Biophys. Acta 173,
501
- Yanofsky C., Lennox E.S. (1959). Virology 8, 425
- Yanofsky C., Carlton B.C., Guest J.R., Helinski D.R., Henning U.
(1964). Proc. Natl. Acad. Sci. U.S. 51, 266
- Yanofsky C., Horn V., Bonner M., Stasiowski S. (1971).
Genetics 69, 409
- Yarmolinski M., Jordan E., Wiesmeyer H. (1961). Cold Spring
Harbor Symp. Quant. Biol. 26, 217
- Zabin I., Kepes A., Monod J. (1959). Biochem. Biophys. Res.
Comm. 1, 289
- Zabin I., Kepes A., Monod J. (1962). J. Biol. Chem. 237, 253
- Zabin I. (1963). J. Biol. Chem. 238, 3300
- Zabin I., Fowler A.V. (1970). dans "The Lactose Operon", Cold
Spring Harbor Lab, p. 27
- Zabin I., Fowler A.V. (1972). J. Biol. Chem. 247, 5432
- Zipser D., Zabell S., Rothman J., Grodzicker T., Wenk M.,
Novitski M. (1970). J. Mol. Biol. 49, 251

